

STIC-Biotech/ChemLib

92653

From: Mehta, Ashwin
Sent: Monday, April 28, 2003 3:57 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

STIC,

Please search the commercial and interference databases for the nucleotide sequences of SEQ ID NOs: 5 and 6 from 09/810,764.

My mail room is 9E12, office 9E07, art unit 1638.

Thank you,
Ashwin Mehta

Ashwin Mehta
United States Patent and Trademark Office
Biotechnology Patent Examiner
703-306-4540

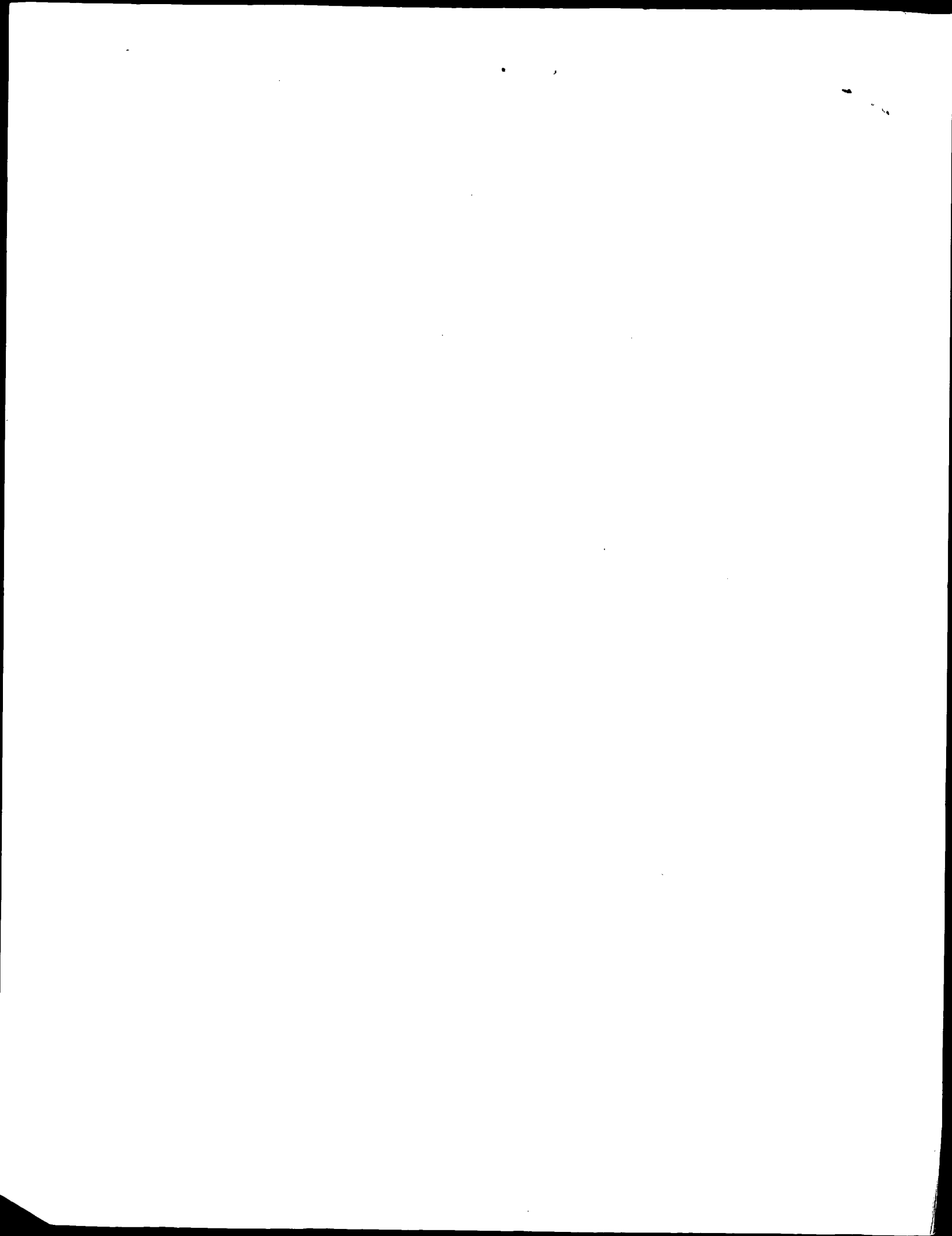
Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-3534

CR1E

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/29
Date Completed: 5/2
Searcher Prep/Review: RV
Clerical: _____
Online time: RV

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:33:59 ; Search time 75 Seconds
(without alignments)
33.757 Million cell updates/sec

Title: US-09-810-764A-5

Perfect score: 90

Sequence: 1 RKGPERKPGVNGTIXXV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	86.7	465	21 AAG40781	Zea mays protein f
2	72	80.0	476	21 AAG18102	Arabidopsis thalia
3	72	80.0	476	21 AAG52762	Arabidopsis thalia
4	72	80.0	497	21 AAG18101	Arabidopsis thalia
5	72	80.0	497	21 AAG52761	Arabidopsis thalia
6	69	76.7	396	22 AAG93183	C glutamicum prote
7	66	73.3	394	21 AAB15890	E. coli proliferat
8	66	73.3	394	22 AAU34769	E. coli cellular p
9	66	73.3	394	22 AAU34838	E. coli cellular p
10	66	73.3	394	22 AAU35466	Haemophilus influe

11	66	73.3	394	22 AAU35476	Haemophilus influe
12	66	73.3	409	22 AAU38371	Salmonella typhi c
13	65	72.2	396	13 AAR20242	Sequence of transl
14	65	72.2	396	13 AAR20243	Sequence of transl
15	65	72.2	396	13 AAR20245	Sequence of elfamy
16	65	72.2	396	13 AAR20246	Sequence of elfamy
17	65	72.2	396	13 AAR20247	Sequence of elfamy
18	65	72.2	396	13 AAR20248	Sequence of elfamy
19	63	70.0	397	23 ABP31625	Human structural p
20	63	70.0	132	22 AAU66514	Propionibacterium
21	63	70.0	436	22 AAU39992	Propionibacterium
22	62	68.9	22	18 AAW33404	N-terminal sequenc
23	62	68.9	22	19 AAW73036	Helicobacter pylor
24	62	68.9	399	19 AAW73035	Helicobacter pylor
25	62	68.9	399	22 AAU35822	Helicobacter pylor
26	61	67.8	397	22 AAU36402	Pseudomonas aerugi
27	61	67.8	397	22 AAU36409	Pseudomonas aerugi
28	60	66.7	395	23 ABB48242	Listeria monocytog
29	59	65.6	396	18 AAW30303	Elongation factor
30	57	63.3	398	22 AAU37578	Streptococcus pneu
31	57	63.3	398	23 ABP30281	Streptococcus poly
32	57	63.3	404	23 ABP25922	Streptococcus poly
33	57	63.3	415	23 ABP25923	Streptococcus poly
34	56	62.2	489	22 ABB62027	Drosophila melanog
35	55	61.1	77	21 AAG12253	Zea mays protein f
36	55	61.1	86	21 AAG12252	Zea mays protein f
37	55	61.1	111	21 AAG12476	Zea mays protein f
38	55	61.1	129	21 AAG12474	Zea mays protein f
39	55	61.1	156	21 AAG44634	Zea mays protein f
40	55	61.1	175	21 AAG44633	Zea mays protein f
41	55	61.1	426	21 AAB03444	Candida albicans e
42	53	58.9	394	20 AAX34672	Chlamydia pneumoni
43	53	58.9	394	22 AAU38889	C. pneumoniae CT32
44	53	58.9	394	22 AAU38909	C. trachomatis CT3
45	53	58.9	394	23 ABB94303	Chlamydia pneumoni

ALIGNMENTS

RESULT 1

AAG40781
ID AAG40781 standard; Protein: 465 AA.

XX AAG40781;

AC AAG40781;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 50648.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.

OS Zea mays subsp. mays.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 23-MAR-1999; 99US-0123548.

XX 25-MAR-1999; 99US-0125788.

XX 29-MAR-1999; 99US-0126264.

XX 01-APR-1999; 99US-0127482.

XX 06-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.


```

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      86.7%; Score 78; DB 21; Length 465;
Best Local Similarity 84.2%; Pred. No. 5.4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNICTIXV 19
   ||||| |||||
Db 59 RGFERTKPGVNICTIGHV 77

RESULT 2
AAG18102
ID AAG18102 standard; Protein; 476 AA.
XX
AC AAG18102;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 19377.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

```

PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147036.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159299.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 476;
 Best Local Similarity 78.9%; Pred. No. 0.00059;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTKPGVNIIGITXXV 19
 ||||| || ||||| |
 Db 70 RGKFKTKPGVNIIGITGHV 88

RESULT 3
 AAG52762
 ID AAG52762 standard; Protein; 476 AA.
 XX
 AC AAG52762;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 67107.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 476;
Best Local Similarity 78.9%; Pred. No. 0.00059;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RKKFKTPGVNIGTIXV 19

||||||| |||||||

Db 70 RKKFKRPPHVNIGTIGHV 88

RESULT 4

AAG18101
ID AAG18101 standard; Protein; 497 AA.

XX AC AAG18101;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 19376.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146393.

```
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157553.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 497;
Best Local Similarity 78.9%; Pred. No. 0.00062;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKPERTKPGVNIGTIXXV 19
   ||||| ||||| |
Db 91 RGKPERKKPHVNIQTIGHV 109

RESULT 5
AAG52761
ID AAG52761 standard; Protein; 497 AA.
XX
AC AAG52761;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67106.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
```

PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 28-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 497;
 Best Local Similarity 78.9%; Pred. No. 0.00062;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXV 19
 Db 91 RGKFKTPGVNIGTIGHV 109

RESULT 6

AAG93183
 ID AAG93183 standard; Protein: 396 AA.

AC AAG93183;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6937.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 OS organic acid synthesis.

OS Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI: 2001-376931/40.

DR N-PSDB; AAH68402.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT

XX Claim 29; SEQ ID NO: 6937; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

SX Sequence 396 AA;

Query Match 76.7%; Score 69; DB 22; Length 396;
 Best Local Similarity 73.7%; Pred. No. 0.0016;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXV 19

Db 3 KAKFKTPGVNIGTIGHV 21

RESULT 7

AAAB15890
 ID AAAB15890 standard; Protein: 394 AA.

XX AAB15890;
 AC 05-OCT-2000 (first entry)
 DT E. coli proliferation associated protein sequence SEQ ID NO:247.
 DE Escherichia coli; E. coli; proliferation; inhibition; screening;
 XX antimicrobial; bacterial growth; antisense therapy; antibacterial.
 KW Escherichia coli.
 OS Escherichia coli.
 XX WO2000044906-A2.
 PN 03-AUG-2000.
 PD 27-JAN-2000; 2000WO-US02200.
 PF 27-JAN-1999; 99US-0117405.
 PR (ELIT-) ELITRA PHARM INC.
 PA Zyskind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI WPI: 2000-514822/46.
 DR N-PSDB; AAAB5894.
 DR Novel polynucleotides and polypeptides associated with microorganism
 PT proliferation, used to identify inhibitors of bacterial growth and
 PT proliferation, for use in antisense therapy -
 XX Claim 11; Page 177-178; 316pp; English.
 PS AAAB5809 to AAAB5889 and AAAB66058 to AAAB66138 represent nucleotide
 CC sequences derived from Escherichia coli which inhibit E. coli
 CC proliferation. AAAB5890 to AAAB66055 and AAAB15886 to AAAB16040 represent
 CC nucleotide and protein sequences associated with E. coli proliferation.
 CC AAAB66056 and AAAB66057 represent primers used for sequencing E. coli
 CC proliferation inhibiting nucleotide inserts in an example from the
 CC present invention. Methods from the present invention can be used to
 CC identify a proliferation- required gene in a microorganism, by contacting
 CC a microorganism with a proliferation-required gene activity inhibitory
 CC nucleic acid identified in another organism, and determining if
 CC inhibition occurs in the second microorganism. The nucleic acid sequences
 CC identified as being required for bacterial growth and proliferation, can
 CC be used for antisense therapy for killing bacteria.
 XX Sequence 394 AA;
 SQ Query Match 73.3%; Score 66; DB 21; Length 394;
 Best Local Similarity 76.5%; Pred. No. 0.0052;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 KFKTKTPGVNIGTIXV 19
 Db 5 KFKTKTPGVNIGTIGHV 21
 RESULT 8
 AAU34769
 ID AAU34769 standard; Protein: 394 AA.
 XX AAU34769;
 AC AAU34769;
 XX 14-FEB-2002 (first entry)
 DT E. coli cellular proliferation protein #350.
 DE Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX Escherichia coli.

XX	
XX	WO200170955-A2.
XX	
XX	27-SEP-2001.
XX	
XX	21-MAR-2001; 2001WO-US09180.
XX	
XX	21-MAR-2000; 2000US-191078P.
XX	23-MAY-2000; 2000US-206848P.
XX	26-MAY-2000; 2000US-207727P.
XX	23-OCT-2000; 2000US-242578P.
XX	27-NOV-2000; 2000US-253625P.
XX	22-DEC-2000; 2000US-257931P.
XX	16-FEB-2001; 2001US-269308P.
XX	(ELIT-) ELITRA PHARM INC.
XX	
XX	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX	Yamamoto RT, Xu HH;
XX	
XX	WPI; 2001-611495/70.
XX	N-PSDB; AAS52628.
XX	
XX	New polynucleotides for the identification and development of
XX	antibiotics, comprise sequences of antisense nucleic acids -
XX	
XX	Example 3; Seq ID No 10362; 511pp; English.
XX	
XX	The invention relates to antisense inhibitors of genes essential to
XX	prokaryotic cellular proliferation, their use in identifying the
XX	genes, their use in the discovery of novel antibiotics, the essential
XX	genes themselves and the encoded proteins. The prokaryotes used are
XX	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX	invention is also useful for the identification of potential new targets
XX	for antibiotic development. The antisense nucleic acids can also be used
XX	to identify proteins used in proliferation, to express these proteins,
XX	and to obtain antibodies capable of binding to the expressed proteins.
XX	The proteins can be used to screen compounds in rational drug discovery
XX	programmes. The antisense nucleic acid sequence is also useful to screen
XX	for homologous nucleic acids which are required for cell proliferation in
XX	a wide variety of organisms. The present sequence represents an
XX	essential prokaryotic cellular proliferation protein.
XX	Note: The sequence data for this patent did not form part
XX	of the printed specification, but was obtained in electronic
XX	format directly from WIPO at
XX	ftp.wipo.int/pub/published pct sequences.
XX	

```

XX
SQ      Sequence      394 AA;
Query Match      73.3%; Score 66; DB 22; Length 394;
Best Local Similarity 76.58; Pred. NO. 0.0052;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      3 KFERTKPGVNIQTIXXV 19
        ||||| ||:| |
        ||||| ||:| |
Db      5 KFERTKPHVNVGTIGHV 21

RESULT 9
AAU34838
ID      AAU34838 standard; Protein; 394 AA.
XX
XX      AAU34838;
XX
XX      14-FEB-2002 (first entry)
DT
DE      E. coli cellular proliferation protein #419.
XX
XX      Antisense; prokaryotic cellular proliferation protein;
KW      antibiotic; antibacterial; drug design.
XX
XX      Escherichia coli.
OS

```



```

XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS53325.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids .
XX PS Example 3; Seq ID No 11059; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 394 AA;

Query Match 73.3%; Score 66; DB 22; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0052;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFEKTRPGVNIQIXXV 19
Db ||||| ||:|:| |
5 KFEKTRPHVNVGTIGHV 21

RESULT 11
AAU35476
ID AAU35476 standard; Protein; 394 AA.
XX AC AAU35476;
XX DT 14-FEB-2002 (first entry)
XX DE Haemophilus influenzae cellular proliferation protein #117.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Haemophilus influenzae.

```

```

XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS53335.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 11069; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 394 AA;

Query Match 73.3%; Score 66; DB 22; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0052;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFEKTRPGVNIQIXXV 19
Db ||||| ||:|:| |
5 KFEKTRPHVNVGTIGHV 21

RESULT 12
AAU38371
ID AAU38371 standard; Protein; 409 AA.
XX AC AAU38371;
XX DT 14-FEB-2002 (first entry)
XX DE Salmonella typhi cellular proliferation protein #262.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Salmonella typhi.

```

XX WO200170955-A2.
 XX 27-SEP-2001.
 XX 21-MAR-2001; 2001WO-US09180.
 XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI: 2001-611495/70.
 DR N-PSDB; AAS56230.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX Example 3; Seq ID No 13964; 511pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 409 AA;
 SQ
 Query Match 73.3%; Score 66; DB 22; Length 409;
 Best Local Similarity 76.5%; Pred. No. 0.0054; 3; Indels 0; Gaps 0;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 KFERTKPGVNIQTIXXV 19
 DB 20 KFERTKPHVNVGTIGHV 36
 RESULT 13
 AAR20242
 ID AAR20242 standard; Protein; 396 AA.
 AC AAR20242;
 XX 15-APR-1992 (first entry)
 DT Sequence of translation elongation factor Tuf1 encoded by tuf1 gene.
 DE Elfamycin resistant actinomycetes; antibiotic resistant;
 KW elongation factor.
 XX Streptomyces ramocissimus.
 OS

XX EP466251-A.
 XX 15-JAN-1992.
 XX 02-JUL-1991; 91EP-0201702.
 XX 02-JUL-1991; 91EP-0201702.
 PR 10-JUL-1990; 90EP-0201851.
 XX (KONN) GIST-BROCADES NV.
 XX Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 PI Woudt LP;
 XX WPI: 1992-017874/03.
 DR N-PSDB; AAQ20215.
 XX New protein conferring resistance to elfamycin - used to
 PT transform streptomycetes to resistant pheno-type
 XX Example; Fig 1 and Pages 13-15; 35pp; English.
 XX Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EF-TuR). The advantage of this change is that
 CC the limiting factor for the prodn. of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein.
 CC resistant to the elfamycin, pref. mocimycin (Kirmomycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.
 XX Sequence 396 AA;
 SQ
 Query Match 72.2%; Score 65; DB 13; Length 396;
 Best Local Similarity 63.2%; Pred. No. 0.0077; 4; Indels 0; Gaps 0;
 Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RGFERTKPGVNIQTIXXV 19
 DB 2 KAFERTKPHVNVGTIGHI 20
 RESULT 14
 AAR20243
 ID AAR20243 standard; Protein; 396 AA.
 XX AAR20243;
 XX 15-APR-1992 (first entry)
 DT Sequence of translation elongation factor Tuf2 encoded by tuf2 gene.
 DE Elfamycin resistant actinomycetes; antibiotic resistant;
 KW elongation factor.
 XX Streptomyces ramocissimus.
 OS
 XX EP466251-A.
 XX 15-JAN-1992.
 XX 02-JUL-1991; 91EP-0201702.
 XX 02-JUL-1991; 91EP-0201702.
 PR 10-JUL-1990; 90EP-0201851.
 XX (KONN) GIST-BROCADES NV.
 XX Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 PI Woudt LP;
 XX WPI: 1992-017874/03.
 DR N-PSDB; AAQ20216.

GenCore version 5.1.4_p5_4578
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:38:19 ; Search time 43 Seconds
(without alignments)
42.478 Million cell updates/sec

Title: US-09-810-764A-5
Perfect score: 90
Sequence: 1 RGKFKTPGVNIGTIXXV 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	82.2	396	2 D60663	translation elonga
2	72	80.0	409	2 AB2348	translation elonga
3	72	80.0	476	2 S09152	translation elonga
4	72	80.0	478	2 JQ2240	translation elonga
5	72	80.0	478	2 S36183	translation elonga
6	72	80.0	479	2 S21567	translation elonga
7	72	80.0	479	2 S60659	translation elonga
8	72	80.0	485	2 S36184	translation elonga
9	70	77.8	18	2 S39153	translation elonga
10	70	77.8	46	2 S23851	translation elonga
11	70	77.8	409	2 S78248	translation elonga
12	69	76.7	391	2 AE3346	protein translatio
13	69	76.7	394	2 C71672	translation elonga
14	69	76.7	394	2 S62726	translation elonga
15	69	76.7	394	2 H97825	elongation factor
16	69	76.7	396	2 S31151	translation elonga
17	69	76.7	396	2 G87143	elongation factor
18	69	76.7	405	2 E75533	translation elonga
19	69	76.7	406	2 AH3344	translation elonga
20	69	76.7	409	1 EFGT	protein translatio
21	69	76.7	488	2 T06821	translation elonga
22	68	75.6	397	2 S23908	translation elonga
23	68	75.6	397	2 S50138	translation elonga
24	68	75.6	397	2 PC4060	translation elonga
25	68	75.6	405	2 C60663	translation elonga
26	68	75.6	409	2 S04430	translation elonga
27	67	74.4	394	2 B60663	translation elonga
28	67	74.4	397	2 H97285	elongation factor
29	67	74.4	410	2 S04391	translation elonga

30 66 73.3 394 1 ERECT
31 66 73.3 394 1 ERECTA
32 66 73.3 394 2 S13561
33 66 73.3 394 2 S13560
34 66 73.3 394 2 E64078
35 66 73.3 394 2 F91152
36 66 73.3 394 2 G91241
37 66 73.3 394 2 D82332
38 66 73.3 394 2 G82337
39 66 73.3 394 2 D86089
40 66 73.3 394 2 AB0457
41 66 73.3 394 2 AE0025
42 66 73.3 394 2 B85998
43 66 73.3 394 2 AD0934
44 66 73.3 394 2 AD1005
45 66 73.3 396 2 A44795

ALIGNMENTS

RESULT 1
D60663

translation elongation factor EF-Tu - Pseudomonas cepacia

C:Species: Pseudomonas cepacia

C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 02-Feb-2001

C:Accession: D60663

R:Ludwig, W.; Weizenegger, M.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moell

Arch. Microbiol. 153, 241-247, 1990

A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongation factor Tu homologues

A:Reference number: A60663; MUID:90240875; PMID:2110445

A:Accession: D60663

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-396 <IUD>

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu

C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:13-139/Domain: translation elongation factor Tu homology <ETU>

F:19-26/Region: nucleotide-binding motif A (P-loop)

F:136-139/Region: GTP-binding NKXD motif

F:174-176/Region: GTP-binding SAK/L motif

F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asp, Ser)

Query Match

Best Local Similarity 82.2%; Score 74; DB 2; Length 396;

Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXXV 19

Db 3 KGKFKTPGVNIGTIXXV 21

RESULT 2
AB2348

translation elongation factor EF-Tu [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AB2348

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2348

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076036.1; PID:g17133473; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: tufa

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu

us-09-810-764a-5.rpr

Fri May 2 08:51:22 2003

C;Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F:1-69/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:70-478/Product: translation elongation factor Tu #status predicted <MAT>
 F:82-208/Domain: translation elongation factor Tu homology <ETU>
 F:88-95/Region: nucleotide-binding motif A (P-loop)
 F:205-208/Region: GTP-binding NKXD motif
 F:243-245/Region: GTP-binding SAK/L motif
 F:94,95,131,205,206,208,243/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser)

Query Match 80.0%; Score 72; DB 2; Length 478;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 ||||| || ||||| |
 Db 72 RGFERTKPGVNIQTIXV 90

RESULT 5

S36183
 translation elongation factor EF-Tu.A precursor, chloroplast - wood tobacco
 C;Species: Nicotiana sylvestris (wood tobacco)
 C;Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 02-Feb-2001
 C;Accession: S36183; S37674; S46548; S40614
 R;Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiyama, M.
 Plant Mol. Biol. 22, 767-774, 1993
 A;Title: Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco
 A;Reference number: S36183; MUID:93363910; PMID:8958028
 A;Accession: S36183
 A;Molecule type: mRNA
 A;Residues: 1-478 <MBR1>
 A;Cross-references: GB:D11375
 R;Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiyama, M.
 submitted to the EMBL Data Library, June 1992
 A;Reference number: S37674
 A;Accession: S37674
 A;Molecule type: mRNA
 A;Residues: 1-457 <MBR2>
 A;Cross-references: GB:D11375; NID:g218309; PIDN:BAA01974.1; PID:g218310
 R;Sugita, M.; Murayama, Y.; Sugiyama, M.
 submitted to the EMBL Data Library, June 1992
 A;Description: Structure and differential expression of two distinct genes encoding
 A;Reference number: S46548
 A;Accession: S46548
 A;Molecule type: DNA
 A;Residues: 1-478 <SUG1>
 A;Cross-references: EMBL:D11469; NID:g459238; PIDN:BAA02027.1; PID:g459239
 R;Sugita, M.; Murayama, Y.; Sugiyama, M.
 Curr. Genet. 25, 164-168, 1994
 A;Title: Structure and differential expression of two distinct genes encoding the ch
 A;Reference number: S40614; MUID:94373864; PMID:8087886
 A;Accession: S40614
 A;Molecule type: DNA
 A;Residues: 1-111 <SUG2>
 A;Cross-references: EMBL:D11469
 C;Genetics:
 A;Gene: tufa
 C;Superfamily: translation elongation factor Tu; translation elongation factor Tu ho
 C;Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthes
 F:1-70/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:71-478/Product: translation elongation factor Tu.A #status predicted <MAT>
 F:82-208/Domain: translation elongation factor Tu homology <ETU>
 F:88-95/Region: nucleotide-binding motif A (P-loop)
 F:205-208/Region: GTP-binding NKXD motif
 F:243-245/Region: GTP-binding SAK/L motif
 F:94,95,131,205,206,208,243/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser)

Query Match 80.0%; Score 72; DB 2; Length 478;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 ||||| || ||||| |
 Db 72 RGFERTKPGVNIQTIXV 90

QY 1 RGFERTKPGVNIQTIXV 19
 ||||| || ||||| |

Query Match 80.0%; Score 72; DB 2; Length 409;
 Best Local Similarity 78.9%; Pred. No. 0.00021;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 ||||| || ||||| |
 Db 3 RGFERTKPGVNIQTIXV 21

RESULT 3

S09152
 translation elongation factor EF-Tu precursor, chloroplast - Arabidopsis thaliana
 N;Alternate names: protein F9f13.10
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 02-Feb-2001
 C;Accession: S09152; T10578
 R;Baldauf, S.L.; Palmer, J.D.
 Nature 344, 262-265, 1990
 A;Title: Evolutionary transfer of the chloroplast tufa gene to the nucleus.
 A;Reference number: S09152; MUID:90190846; PMID:2314461
 A;Accession: S09152
 A;Molecule type: DNA
 A;Residues: 1-476 <BAL>
 A;Cross-references: EMBL:X52256; NID:g22564; PIDN:CAA36498.1; PID:g22565
 R;Bevan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
 submitted to the Protein Sequence Database, June 1999
 A;Reference number: Z16991
 A;Accession: T10578
 A;Molecule type: DNA
 A;Residues: 15-476 <BEV>
 A;Cross-references: EMBL:AL080253; GSPDB:GN00062; ATSP:F9f13.10
 A;Experimental source: cultivar Columbia; BAC clone F9f13
 C;Genetics:
 A;Gene: tufa; ATSP:F9f13.10
 A;Map position: 4
 A;Genome: nuclear
 C;Superfamily: translation elongation factor Tu; translation elongation factor Tu homol
 C;Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F:1-67/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:68-476/Product: translation elongation factor Tu #status predicted <MAT>
 F:80-206/Domain: translation elongation factor Tu homology <ETU>
 F:86-93/Region: nucleotide-binding motif A (P-loop)
 F:203-206/Region: GTP-binding NKXD motif
 F:241-243/Region: GTP-binding SAK/L motif
 F:92,93,129,203,204,206,241/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #st

Query Match 80.0%; Score 72; DB 2; Length 476;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 ||||| || ||||| |
 Db 70 RGFERTKPGVNIQTIXV 88

RESULT 4

JQ2240
 translation elongation factor EF-Tu precursor - common tobacco chloroplast
 C;Species: chloroplast Nicotiana tabacum (common tobacco)
 C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Feb-2001
 C;Accession: JQ2240
 R;Ursin, V.M.; Becker, C.K.; Shewmaker, C.K.
 Plant Physiol. 101, 333-334, 1993
 A;Title: Cloning and nucleotide sequence of a tobacco chloroplast translational elongat
 A;Reference number: JQ2240; MUID:94105295; PMID:8278900
 A;Accession: JQ2240
 A;Molecule type: DNA
 A;Residues: 1-478 <URS>
 C;Comment: This protein is an essential component for protein synthesis that functions
 C;Genetics:
 A;Genome: chloroplast
 C;Superfamily: translation elongation factor Tu; translation elongation factor Tu homol

Query Match 80.0%; Score 72; DB 2; Length 476;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 ||||| || ||||| |
 Db 70 RGFERTKPGVNIQTIXV 88

Db 72 RGKFERKPPHVNIGTIGHV 90

RESULT 6

S21567

translation elongation factor EF-Tu precursor - soybean chloroplast

C:Species: chloroplast Glycine max (soybean)

C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001

C:Accession: S21567

R:Bonny C.; Stutz, E.

submitted to the EMBL Data Library, May 1992

A:Description: Soybean (glycine max l.) nuclear DNA: structure and expression of chlorop

A:Reference number: S21567

A:Accession: S21567

A:Molecule type: DNA

A:Residues: 1-479 <HON>

A:Cross-references: EMBL:X66062; NID:g18775; PIDN:CAA46864.1; PID:g18776

C:Genetics:

A:Gene: tuftA

A:Genome: chloroplast

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:83-209/Domain: translation elongation factor Tu homology <ETU>

F:89-96/Region: nucleotide-binding motif A (P-loop)

F:206-209/Region: GTP-binding NKXD motif

F:244-246/Region: GTP-binding SAK/L motif

F:95,96,132,206,207,209,244/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #st

Query Match

80.0%; Score 72; DB 2; Length 479;

Best Local Similarity 78.9%; Pred. No. 0.00025;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFERKPPHVNIGTIGHV 91

||||| || ||||| |

Db 73 RGKFERKPPHVNIGTIGHV 91

RESULT 7

S60659

translation elongation factor EF-Tu precursor, chloroplast - soybean

C:Species: Glycine max (soybean)

C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 02-Feb-2001

C:Accession: S60659

R:Maurer, F.; Murone, M.; Stutz, E.

submitted to the EMBL Data Library, July 1995

A:Description: The nuclear genome of soybean contains four tuf genes belonging to two su

A:Reference number: S60659

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-479 <NAU>

A:Cross-references: EMBL:X89058; NID:g949872; PIDN:CAA61444.1; PID:g949873

C:Genetics:

A:Genome: nuclear

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:83-209/Domain: translation elongation factor Tu homology <ETU>

F:89-96/Region: nucleotide-binding motif A (P-loop)

F:206-209/Region: GTP-binding NKXD motif

F:244-246/Region: GTP-binding SAK/L motif

F:95,96,132,206,207,209,244/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #st

Query Match

80.0%; Score 72; DB 2; Length 479;

Best Local Similarity 78.9%; Pred. No. 0.00025;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFERKPPHVNIGTIGHV 91

||||| || ||||| |

Db 73 RGKFERKPPHVNIGTIGHV 91

RESULT 8

S36184

translation elongation factor EF-Tu.B precursor, chloroplast - wood tobacco

C:Species: Nicotiana sylvestris (wood tobacco)

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 02-Feb-2001

C:Accession: S36184; S37675; S46549; S40615

R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiura, M.

Plant Mol. Biol. 22, 767-774, 1993

A:Title: Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco

A:Reference number: S36183; MUID:93363910; PMID:8358028

A:Accession: S36184

A:Molecule type: mRNA

A:Residues: 1-485 <MUR>

A:Cross-references: GB:D11376

R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiura, M.

submitted to the EMBL Data Library, June 1992

A:Reference number: S37674

A:Accession: S37675

A:Molecule type: mRNA

A:Residues: 61-430, 'IMS', 434-448, 'NM', 451-482, 'ILE', 485 <MUR>

A:Cross-references: GB:D11376; NID:g218311; PIDN:BAA01975.1; PID:g218312

R:Sugita, M.; Murayama, Y.; Sugiura, M.

submitted to the EMBL Data Library, June 1992

A:Description: Structure and differential expression of two distinct genes encoding the

A:Reference number: S46548

A:Accession: S46549

A:Molecule type: mRNA

A:Residues: 1-430, 'IMS', 434-448, 'NM', 451-482, 'IL', 485 <SUGI>

A:Cross-references: EMBL:D11470; NID:g459240; PIDN:BAA02028.1; PID:g459241

R:Sugita, M.; Murayama, Y.; Sugiura, M.

Curr. Genet. 25, 164-168, 1994

A:Title: Structure and differential expression of two distinct genes encoding the ch

A:Reference number: S40614; MUID:94373864; PMID:8087886

A:Accession: S40615

A:Molecule type: DNA

A:Residues: 1-103 <SUG2>

C:Genetics:

A:Gene: tufB

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthes

F:1-77/Domain: transit peptide (chloroplast) #status predicted <TNP>

F:78-485/Product: translation elongation factor Tu.B #status predicted <MAT>

F:89-215/Domain: translation elongation factor Tu.B #status predicted <MAT>

F:95-102/Region: nucleotide-binding motif A (P-loop)

F:212-215/Region: GTP-binding NKXD motif

F:250-252/Region: GTP-binding SAK/L motif

F:101,102,138,212,213,215,250/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser

Query Match

80.0%; Score 72; DB 2; Length 485;

Best Local Similarity 78.9%; Pred. No. 0.00025;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFERKPPHVNIGTIXV 19

||||| || ||||| |

Db 79 RGKFERKPPHVNIGTIGHV 97

RESULT 9

S39153

translation elongation factor EF-Tu, chloroplast - common tobacco (fragment)

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 05-Dec-1997

C:Accession: S39153

R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiura, M.

Plant Mol. Biol. 22, 767-774, 1993

A:Title: Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco

A:Reference number: S36183; MUID:93363910; PMID:8358028

A:Accession: S39153

A:Molecule type: protein

A:Residues: 1-18 <MUR>

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

C:Keywords: chloroplast; GTP binding; protein biosynthesis

Query Match

77.8%; Score 70; DB 2; Length 18;

Best Local Similarity 87.5%; Pred. No. 1.9e-05;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 16
| | | | | | | | | |
DB 2 RGFERTKPGVNIQT 17

RESULT 10

S23851
translation elongation factor EF-Tu - Eikenella corrodens (fragment)
C:Species: Eikenella corrodens
C:Date: 12-Feb-1993 #sequence_revision 08-Nov-1996 #text_change 02-Feb-2001
C:Accession: S23851
R: Rao, V.K.; Whitlock, J.A.; Progulskie-Fox, A.
Submitted to the EMBL Data Library, June 1992
A:Description: Cloning, characterization, and sequencing of two hemagglutinin genes from
A:Reference number: S23847
A:Accession: S23851
A:Molecule type: DNA
A:Residues: 1-46 <RAO>
A:Cross-references: EMBL:Z12610; NID:g41653; PIDN:CAA78258.1; PID:g41658
A:Experimental source: Eikenella corrodens ATCC 23834
C:Genetics:
A:Gene: tuftA
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
F:19-26/Region: nucleotide-binding motif A (P-loop)

Query Match 77.8%; Score 70; DB 2; Length 46;

Best Local Similarity 68.4%; Pred. No. 5e-05;

Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
: | | | | | | | | | |
DB 3 RGFERTKPGVNIQT 21

RESULT 11

S78248
translation elongation factor EF-Tu - Odontella sinensis chloroplast
C:Species: chloroplast Odontella sinensis
C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 02-Feb-2001
C:Accession: S78248
R: Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis
A:Reference number: S78238
A:Accession: S78248
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-409 <KOW>
A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91621.1; PID:g1185138
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C:Genetics:
A:Gene: tuftA
A:Genome: Chloroplast
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
F:13-139/Domain: translation elongation factor Tu homology <ETU>
F:19-26/Region: nucleotide-binding motif A (P-loop)
F:136-139/Region: GTP-binding NKXD motif
F:174-176/Region: GTP-binding SAK/L motif
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 77.8%; Score 70; DB 2; Length 409;

Best Local Similarity 78.9%; Pred. No. 0.00046;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
| | | | | | | | | |
DB 3 RGFERTKPGVNIQT 21

RESULT 12

AE3346
protein translation elongation factor Tu (EF-tu) [imported] - Brucella melitensis (st
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AE3346
R: DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanov
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3346
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL51936.1; PID:g17982694; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0755
A:Map position: 1
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 76.7%; Score 69; DB 2; Length 391;

Best Local Similarity 73.7%; Pred. No. 0.00065;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
: | | | | | | | | | |
DB 3 RGFERTKPGVNIQT 21

RESULT 13

C71672
translation elongation factor EF-Tu (tuf) RP661 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 02-Feb-2001
C:Accession: C71672
R: Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: C71672
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-394 <AND>
A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15101.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: tuf; RP661
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 76.7%; Score 69; DB 2; Length 394;

Best Local Similarity 73.7%; Pred. No. 0.00066;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
: | | | | | | | | | |
DB 3 RGFERTKPGVNIQT 21

RESULT 14

S62726
translation elongation factor tu (EF-tu) UU522 [similarity] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
C:Accession: S62726; F82879
R: Kamla, V.

submitted to the EMBL Data Library, June 1994

A:Reference number: S62726
A:Accession: S62726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <RAM>
A:Cross-references: EMBL:Z34275; NID:g498790; PIDN:CAA84029.1; PID:g498791
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir
A:Reference number: A82870
A:Accession: F82879
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <GLA>
A:Cross-references: GB:AE002151; GB:AF222894; NID:g6899524; PIDN:AAF30935.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: tuf; UU522
A:Genetic code: SGC3
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo
C:Keywords: GTP binding; nucleotide binding; P-loop
F:13-139/Domain: translation elongation factor Tu homology <ETU>
F:19-26/Region: nucleotide-binding motif A (P-loop)
F:136-139/Region: GTP-binding NKXD motif

Query Match 76.7%; Score 69; DB 2; Length 394;

Best Local Similarity 73.7%; Pred. No. 0.00066;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGRFERTKPGVNIQTIXV 19

: ||||| |||||

Db 3 KAKFERTKPHVNIQTICHV 21

RESULT 15

H97825

elongation factor EF-Tu [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001

C:Accession: H97825

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R

Science 293; 2093-2098; 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: H97825

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-394 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL03546.1; PID:g15620123; GSPDB:GN00173

C:Genetics:

A:Gene: tuf

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

Query Match

Best Local Similarity 76.7%; Score 69; DB 2; Length 394;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGRFERTKPGVNIQTIXV 19

: ||||| |||||

Db 3 KAKFERTKPHVNIQTICHV 21

Search completed: April 29, 2003, 09:42:20

Job time : 44 secs

OM protein - protein search, using sw model
Run on: April 29, 2003, 09:36:24 ; Search time 25 Seconds
(without alignments)
31.522 Million cell updates/sec

Title: US-09-810-764A-5
Perfect score: 90
Sequence: 1 RGKERTKPGVNICTIXV 19
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	83.3	391	1 EFTU_RHILO	098167 rhizobium l
2	74	82.2	396	1 EFTU_BURCE	P31677 burkholderi
3	72	80.0	476	1 EFTU_ARATH	P17745 arabidopsis
4	72	80.0	478	1 EFTU_TOBAC	P41342 nicotiana t
5	72	80.0	479	1 EFTU_SOYBN	Q43467 glycine max
6	72	80.0	479	1 EFTU_EIKCO	P46280 glycine max
7	70	77.8	46	1 EFTU_ODOSI	P35644 eikenella c
8	70	77.8	409	1 EFTU_ODOSI	Q91462 odontella s
9	69	76.7	394	1 EFTU_NEPOL	P48865 rickettsia
10	69	76.7	394	1 EFTU_UREPA	P50068 ureaplasma
11	69	76.7	395	1 EFTU_BACST	O50306 bacillus st
12	69	76.7	396	1 EFTU_MYCLE	P42439 corynebacte
13	69	76.7	397	1 EFTU_STRCJ	P30768 mycobacteri
14	69	76.7	405	1 EFTU_DEIRA	P95724 streptomyce
15	69	76.7	488	1 EFTU_EUGGR	Q91342 deinococcu
16	69	76.7	396	1 EFTU_PEA	P02991 euglena gra
17	69	76.7	397	1 EFTU_THICU	O24310 pismus sativ
18	69	76.7	397	1 EFTU_STRCO	P42481 thioacillu
19	68	75.6	397	1 EFTI_STRCU	Q40174 streptomyce
20	68	75.6	397	1 EFTI_STRCU	Q53871 streptomyce
21	68	75.6	397	1 EFTI_STRCU	Q33594 streptomyce
22	68	75.6	405	1 EFTU_DEISP	P33168 deionema s
23	68	75.6	409	1 EFTU_SYNTP	P18668 synchococc
24	68	75.6	409	1 EFTU_SYNTP	P33171 synchococc
25	67	74.4	394	1 EFTU_BACFR	P19457 guillardia
26	67	74.4	408	1 EFTU_GUITH	P13552 spirulina p
27	67	74.4	410	1 EFTU_GUITH	P02990 escherichia
28	66	73.3	393	1 EFTU_ECOLI	P43926 haemophilus
29	66	73.3	393	1 EFTU_HAEIN	P21694 salmonella
30	66	73.3	393	1 EFTU_SALTY	P57939 pasteurella
31	66	73.3	394	1 EFTA_PASMU	
32	66	73.3	394	1 EFTA_PASMU	
33	66	73.3	394	1 EFTA_PASMU	

34	66	73.3	394	1	EFTU_PASMU
35	66	73.3	396	1	EFTU_MYCTU
36	66	73.3	409	1	EFTU_PORPU
37	66	73.3	418	1	EFTU_CHLRE
38	65	72.2	397	1	EFTU_CHLRE
39	64	71.1	391	1	EFTU_STRRA
40	64	71.1	391	1	EFTU_STRRA
41	64	71.1	399	1	EFTU_BRELN
42	63	70.0	405	1	EFTU_SYNY3
43	63	70.0	409	1	EFTU_THETH
44	63	70.0	410	1	EFTU_ASTLO
45	62	68.9	394	1	EFTU_MESVI
					EFTU_NEIGO

ALIGNMENTS

RESULT 1
EFTU_RHILO
ID EFTU_RHILO STANDARD; PRT; 391 AA.
AC Q981F7;
DT 16-OCT-2001 (Rel. 40, Created)
DE 16-OCT-2001 (Rel. 40, Last sequence update)
GN Elongation factor Tu (EF-Tu),
OS (TUPA OR MLR0263) AND (TUFB OR MLR0288).
OC Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Watanabe A., Igesawa K., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kishida Y., Kiyokawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Mochizuki Y., Nakayama C., Kobara M., Matsumoto M., Matsuno A.,
RA Takeuchi C., Yamada M., Nakazaki N., Shimpō S., Sugimoto M.,
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti";
RL DNA Res 7:331-338(2000).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILILARY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC use by non-profit institutions. There are no restrictions on its
CC modified and this statement as long as its content is in no way
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP002994; BAB47886.1;
CC EMBL; AP002994; BAB47904.1;
CC InterPro; IPR004541; EF-Tu.
CC InterPro; IPR004160; EFTU_Cterm.
CC InterPro; IPR004161; EFTU_D2.
CC InterPro; IPR000795; EF-GTPbind.
CC Pfam; PF00009; SmallGTP.
CC Pfam; PF03143; GTP_EFTU; 1.
CC PRINTS; PR00315; ELONGATNCT.
CC TIGRfams; TIGR00231; small_GTP; 1.
CC TIGRfams; TIGR00485; EF-Tu; 1.
CC PROSITE; PS00301; EFACOR_GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
KW

us-09-810-764a-5.rsp

Fri May 2 08:51:22 2003

```

RESULT 3
EFTU_ARATH STANDARD; PRT; 476 AA.
ID EFTU_ARATH
AC P17745;
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation precursor (EF-Tu)).
DE Elongation factor Tu, chloroplast precursor (EF-Tu).
DE TUPA OR A14G20360 OR F9F13.10.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Arabidopsidaceae; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Arabidopsid.
OC euroids II; Brassicales; Brassicaceae; Arabidopsid.
OC NCBI_TaxID=3702;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Landsberg erecta;
RC STRAIN=cv. Landsberg erecta; PubMed=2314461;
RX MEDLINE=90190846; PubMed=2314461;
RX Baldauf S.L., Palmer J.D.; the chloroplast tufa gene to the nucleus."
RT "Evolutionary transfer of the chloroplast tufa gene to the nucleus."
RL Nature 344:262-265(1990).
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia; PubMed=10617198;
RX MEDLINE=20083488; PubMed=10617198;
RX Mayer K.F.X., Schueller A., Stiekema W., Entian K.-D., Terry N.,
RX Pohl T., Duesterhoeft A., Brandt P., Grivell L.A., Rieger M., Mueller M.,
RX Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M., Mueller M.,
RX Weichselgartner M., de Simone V., Obermaier B., Mache R., Schmidheini T.,
RX Kreis M., Delsen M., Puigdomenech P., Watson M., Bancroft I.,
RX Reichert B., Portetelie D., Perez-Alonso M., Boutry M., Ridley P.,
RX Vos P., S.-A., McCullagh B., Bilham L., Robben J., Vandenbussche F.,
RX Langham S.-A., McCullagh B., Bilham L., Robben J., Vandenbussche F.,
RX Van der Meer S., Veltjens I., Voort M., Bastiaens I., Aert R., Defoor E.,
RX Braeken M., Wiltjens I., Voort M., Bastiaens I., Aert R., Defoor E.,
RX Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Dirke W.,
RX Holzer B., Brandt A., Peters S., van Staveren M., Koetter P.,
RX Moelkenhoff P., Klein Lankhorst R., Rose M., Hauf J., Van den Daele H.,
RX Moelkenhoff P., Klein Lankhorst R., Rose M., Hauf J., Van den Daele H.,
RX Berner S., Hempel S., Feldpausch M., Lamberth S., de Clercq R.,
RX De Keyser A., Buyschaert C., Gelin A., Quail M., Bray-Allen S.,
RX Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RX Clark L., Doggett J., Hall S., Kay M., Lennard N., McClay K., Mayes R.,
RX Borkova D., Bloeker H., Scharfe M., Grimm M., Benes V., Rehnert T.-H.,
RX Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RX Gabel C., Fuchs M., Fartmann B., Grandjean K., Dauner D., Herzi A.,
RX Neumann S., Argirou F., Clabaud G., Mendenhall A., Feilner R.,
RX Massenet O., Quigley F., Schmidt W., Lecharny A., Aubourg S.,
RX Schnabl S., Hiller R., Berger C., Monfort A., M. Terol J., Torres A.,
RX Chefor F., Cooke R., Vandenberg M., Barges A., Johnson S., Tacon D., Jesse T.,
RX Gibbons T., Weber N., Vandenberg M., Barges A., Johnson S., Tacon D., Jesse T.,
RX Heijnen D., Haase D., Lemcke K., Mewes H.-W., de la Bastide M., Habermann K.,
RX Frishman D., Schwarz S., Scholler P., Scholler P., Stocker S.,
RX Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RX Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RX Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RX Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RX Latreille B., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RX Minx P., Bentley D., Fulton B., Miller N., Pepin K., Hillier L.,
RX Kramer J., Fulton L., Ryan E., Mardis E., Dante M., Geisel C., Joshua C.,
RX Du H., Ali J., Berghoff A., Jones K., Sun H., Lanar B., Jordan C.,
RX Antoniou B., Zidanic M., Strong C., Shekher M., Matero A., Shah R.,
RX Ma P., Zhong J., Preston R., Vil D., Shekher M., Hoffman J., Till S.,
RX Granat S., Shohdy N., Hasegawa A., Rodriguez M., Hoffman J., Till S.,
RX Chen E., Marra M., Martienssen R., McCombie W.R.;
RN "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RT Nature 402:769-777(1999).
RL -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC -1- AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC

```

0;

Gaps

0;

Length 396;

DB 1;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

Score 74;

Pred. No. 1.6e-05;

Indels

3;

Mismatches

2;

CC BIOSYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-TU/EF-1A SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52356; CAA36498.1; -
 CC EMBL; AL080253; CAB45802.2; -
 CC EMBL; AL161552; CAB79036.1; -
 CC PIR; S09152; S09152.
 CC HSSP; P02990; LEFU.
 CC SWISS-2DPAGE; P17745; ARATH.
 CC InterPro; IPR004541; EF-Tu.
 CC InterPro; IPR004160; EFTU_Cterm.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; EF-GTPbind.
 CC InterPro; IPR005225; Small_GTP.
 CC Pfam; PF00009; GTP-EFTU; 1.
 CC Pfam; PF03143; GTP-EFTU_D3; 1.
 CC Pfam; PF03144; GTP-EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFC.
 CC TIGRFAMS; TIGR00231; small_GTP; 1.
 CC TIGRFAMS; TIGR00485; EF-Tu; 1.
 CC PROSITE; PS00301; EFACOR_GTP; 1.
 CC Elongation factor; Protein biosynthesis; GTP-binding;
 CC Transit peptide; Chloroplast.
 CC TRANSIT 1 67
 CC CHAIN 68 476 CHLOROPLAST (POTENTIAL).
 CC NP_BIND 86 93 GTP (BY SIMILARITY).
 CC NP_BIND 148 152 GTP (BY SIMILARITY).
 CC NP_BIND 203 206 GTP (BY SIMILARITY).
 CC SEQUENCE 476 AA; 51630 MW; 6041F540224A1738 CRC64;
 CC
 CC Query Match 80.0%; Score 72; DB 1; Length 476;
 CC Best Local Similarity 78.9%; Pred. No. 4.3e-05;
 CC Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 RKGFEKTKPGVNIQTIXXV 19
 CC ||||| || ||||| |
 CC Db 70 RKGFEKTKPGVNIQTIXXV 88
 CC
 CC RESULT 4
 CC EFTU_TOBAC STANDARD; PRT; 478 AA.
 CC AC P41342;
 CC DT 01-FEB-1995 (Rel. 31, Created)
 CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
 CC DE 15-DEC-1998 (Rel. 37, Last annotation update)
 CC GN TUFA.
 CC OS Nicotiana tabacum (Common tobacco), and
 CC OS Nicotiana sylvestris (Wood tobacco).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 CC OX NCBI_TaxID=4097, 4096;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC SPECIES-N.Tabacum; STRAIN=cv. NK 236;
 CC RA MEDLINE=94105295; PubMed=8278500;
 CC RX Ursin V.M., Becker C.K., Shewmaker C.K.;
 CC RT "Cloning and nucleotide sequence of a tobacco chloroplast
 CC translational elongation factor, EF-Tu.";
 CC RL plant Physiol. 101:333-334(1993).
 CC [2]

RP SEQUENCE FROM N.A.
 RC SPECIES-N.sylvestris;
 RX MEDLINE=94373864; PubMed=8087886;
 RA Sugita M., Murayama Y., Sugiyama M.;
 RT "Structure and differential expression of two distinct genes encoding
 RT the chloroplast elongation factor Tu in tobacco.";
 RL Curr. Genet. 25:164-168(1994).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
 CC BIOSYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-TU/EF-1A SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M94204; AAA18546.1; -
 CC EMBL; D11469; BAA02027.1; -
 CC PIR; JQ2240; JQ2240.
 CC PIR; S46548; S46548.
 CC HSSP; P02990; LEFU.
 CC InterPro; IPR004541; EF-Tu.
 CC InterPro; IPR004160; EFTU_Cterm.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; EF-GTPbind.
 CC InterPro; IPR005225; Small_GTP.
 CC Pfam; PF00009; GTP-EFTU; 1.
 CC Pfam; PF03143; GTP-EFTU_D3; 1.
 CC Pfam; PF03144; GTP-EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFC.
 CC TIGRFAMS; TIGR00231; small_GTP; 1.
 CC TIGRFAMS; TIGR00485; EF-Tu; 1.
 CC PROSITE; PS00301; EFACOR_GTP; 1.
 CC Elongation factor; Protein biosynthesis; GTP-binding;
 CC Transit peptide; Chloroplast.
 CC TRANSIT 1 69
 CC CHAIN 70 478 CHLOROPLAST (POTENTIAL).
 CC NP_BIND 88 95 GTP (BY SIMILARITY).
 CC NP_BIND 150 154 GTP (BY SIMILARITY).
 CC NP_BIND 205 208 GTP (BY SIMILARITY).
 CC SEQUENCE 478 AA; 51956 MW; 9811609E5FAC001 CRC64;
 CC
 CC Query Match 80.0%; Score 72; DB 1; Length 478;
 CC Best Local Similarity 78.9%; Pred. No. 4.3e-05;
 CC Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 RKGFEKTKPGVNIQTIXXV 19
 CC ||||| || ||||| |
 CC Db 72 RKGFEKTKPGVNIQTIXXV 90
 CC
 CC RESULT 5
 CC EFTU_SOYBN STANDARD; PRT; 479 AA.
 CC AC Q43467;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DE 16-OCT-2001 (Rel. 40, Last annotation update)
 CC GN TUFA.
 CC OS Glycine max (Soybean).
 CC OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 CC OX NCBI_TaxID=3847;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.

the Swiss Institute. Copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation. The EMBL outstation is a non-profit institution and the EMBL outstation is not a company. This statement is as long as its content is not modified and this statement is not removed. Usage by and in non-commercial purposes is allowed. Please send an email to license@ebi.ac.uk for more information. (See http://www.ebi.ac.uk/comm/embldb/212610_CAA784258.1).

```

atch      9FFAE6791379235 CRC64;
          77.8%; Score 70; DB 1; Length 46;
          Similarity 66.4%; Pred No. 8.2e-06;
          13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
          RCKFERTPGVNGTGTXXV 19
          :|||||:| |:|||||
          KGGKERSKPHNVNVTGTCU

```

1996 (Rel. 33, Created)
1996 (Rel. 33, Last Sequence update)
1998 (Rel. 37, Last annotation update)
ion factor Tu (EF-Tu).
ta sinensis (Marine centric diatom).
last.
a; stramen-

ACIL-TRNA TO THE A-SITE OF THE GTP-DEPENDENT BINDING OF
 CELLULAR RIBOSOMES DURING PROTEIN
 SYNTHESIS.
 LOCALIZATION: Chloroplast.
 ACTIVITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY
 EF-1A SUBFAMILY.
 PROTEIN entry is complete.

—0000; I.

CC CC CC CC CC CC DR DR DR

Query	Best	Match	Qy	Db
1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25
26	27	28	29	30
31	32	33	34	35
36	37	38	39	40
41	42	43	44	45
46	47	48	49	50
51	52	53	54	55
56	57	58	59	60
61	62	63	64	65
66	67	68	69	70
71	72	73	74	75
76	77	78	79	80
81	82	83	84	85
86	87	88	89	90
91	92	93	94	95
96	97	98	99	100

DT	01-FH	01-FH	15-DE	Elong	TUFA.	Odont	Chlor	Eukary	Bida
DT	DT	DT	DE	GN	QS	OG	OC	OC	OC

CC
CC
CC
CC
CC
CC
CC
CC

RESULT 9
EFTU_NEPOL
D EFTU_NEPOL
C Q9TKZS;
F 15-JUN-2002
15-JUN-2002
STANDARD;
PRT; 410 AA.
(Rel. 41. Crea...)

Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
Chlorodendreales; Chlorodendraceae; Nephroleimn.
[1]
SEQUENCE FROM N. A.
STRAIN-NIES-484.
MEDLINE-99398694; PubMed-10468594.
Turnell M., Otis C.

his SWISS-PROT entry is correct between the GTP-BINDING ELONGATION FACTOR FAMILY.

EFPRO; IPR004341; EF-Tu.
TS; PRO0315; EFU_D2; 1.
EFPRO; IPR004160; EFU_U_Cterm.
EFPRO; IPR004161; EFUT_D2.
EFPRO; IPR000795; EFCT_D2.
EFPRO; IPR0005225; Small_GTP.
PF00069; Small_GTP.
PF03143; GTP-EFTu; 1-GTP.
PF03144; GTP-EFTu_D3; 1.
TS; PRO0315; EFU_D2; 1.
FAMS; TIGR00231; ELONGATNFACT.
FAMS; TIGR00485; Small_GTP.
FAMS; TIGR00485; Small_GTP; 1.

DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR000315; ELONGATNFACT; 1.
DR TIGRFS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00485; EF-Tu_1
DR PROSITE; PS00301; EFACITOR_GNP; 1.
DR PROSITE; PS00301; EFACITOR_GNP; 1.
DR Elongation factor; Protein biosynthesis: GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
FT NP_BIND 198 198 S->T (IN REF. 1).
FT CONFLICT 205 207 RAT -> ELQ (IN REF. CRC64;
FT CONFLICT 205 207 510FEB2997E389BBF
FT SEQUENCE 394 AA; 42921 MW; 1 Length 394;
SQ

Query Match Similarity 76.7%; Score 69; DB 1; Length 394;
Best Local Similarity 73.7%; Pred. No. 0.00012; Indels 0; Gaps
Matches 14; Conservative 1; Mismatches 4;

QY 1 RGKFKRTPGVNGITGXXV 19 PRT: 394 AA.
DB 3 KAKFKTKPHVNICTIGHV 21

RESULT 11 STANDARD:
ID EFTU_UREPA
IC P50068.1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Long-tail factor Tu (EF-Tu).
DE TU OF OR U0522.
GN Ureaplasma parvum (Ureaplasma urealyticum biotype 1), Ureaplasma.
OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OC NCBI_TaxID=134821;
OX NCBI_TaxID=134821;
RN SEQUENCE FROM N.A.;
RP STRAIN=ATCC 33697 / Serovar 14;
RA Bruck A.;
RL Thesis (1994), Heinrich-Heine University / Duesseeldorf, Germany.
RN SEQUENCE FROM N.A.
RP STRAIN=Serovar 3; PubMed=11048724;
RX MEDLINE=20500219; Glass J.S., Heiner C.R., Chen E.Y.,
RX Glasell J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RX Glasell G.H.;
RA Cape complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum." 762(2000).
RT Nature 407:757-762(2000).
RL FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC SUBUNIT: MONOMER (BY SIMILARITY).
CC SUBCELLULAR LOCATION: CYTOPLASMIC.
CC SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR F
CC EF-TU/EF-I_A SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a
CC between the Swiss Institute of Bioinformatics and the EMBL
CC the European Bioinformatics institutions as long as its content is
CC use by non-profit institutions is not removed.
CC modified and this statement is not removed.
CC modified requires a license agreement (see http://www.isb-sib.ch).
CC or send an email to license@isb-sib.ch).

PROSITE; PS00301; EFACOR_GTP: 1. Chloroplast;
Elongation factor; Protein biosynthesis; Chloroplast;
GTP (BY SIMILARITY).
NP_BINDING 19 26 GTP (BY SIMILARITY).
NP_BIND 81 85 GTP (BY SIMILARITY).
NP_BIND 136 139 GTP (BY SIMILARITY).
NP_BIND 410 AA; 44810 MW; A7A9B0A6E99C1B23 CRC64;
SEQUENCE 410 AA; 44810 MW; A7A9B0A6E99C1B23 CRC64;
77.8%; Score 70; DB 1; Length 410;
Query Match Pred No. 8.3e-05; Indels 0; Gaps 0;
Best Local Similarity 78.9%; Mismatches 4;
Matches
1 RGKFKTKPGVNIQTIXV 19
| | | | | | | | | | | | | | | | | | | | | |
3 RKFKTKPHVNIQTIGHV 21
DB

RESULT 10
EFU_RICPR STANDARD: PRT: 394 AA.
ED EFTU_RICPR
AC P48865; 1996 (Rel. 33, Created)
DC 01-FEB-1996 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 40, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT Elongation factor Tu (EF-Tu).
DE TUF OR RP61.
GN Rickettsia prowazekii.
OS Rickettsia proteobacteria; alpha subdivision; Rickettsiales.
OC Bacteriata; Proteobacteria; Rickettsiales; Rickettsia.
OC Rickettsiales; Rickettsiales; Rickettsia.
OC NCBI_taxid=782;
OX OX
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=97047977; PubMed=892818;
RX Syvanen A., Amiri H., Jamal A., Andersson S.G.E., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E;
RC MEDLINE=99039499; PubMed=9823893;
RX Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA Sichenitz-Ponten T., Alsmark U.C.E., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of mitochondria".
RT Nature 396:133-140 (1998).
RL J. Bacteriol. 178:6192-6199 (1996).
RL [2]
RN


```

DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATNFT.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 394 AA; 42902 MW; 80A887B6C59883E0 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 394;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
Db 3 KAKFERTKPHVNIQTIGHV 21

RESULT 12
EFTU_BACST
ID EFTU_BACST STANDARD; PRT; 395 AA.
AC O50306;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 2184;
RX MEDLINE=98443240; PubMed=9769211;
RA Krasny L., Mesters J.R., Tieleman L.N., Kraal B., Fucik V.,
RA Hilgenfeld R., Jonak J.;
RT "Structure and expression of elongation factor Tu from Bacillus
RL stearothermophilus.";
RL J. Mol. Biol. 283:371-381(1998).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ000260; CAA03976.1;
CC HSP: P02990; IETU.
CC InterPro: IPR004541; EF-Tu.
CC InterPro: IPR004160; EFTU_Cterm.
CC InterPro: IPR004161; EFTU_D2.
CC InterPro: IPR00795; EF_GTPbind.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03143; GTP_EFTU_D3; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 1.
CC PRINTS: PR00315; ELONGATNFT.
CC TIGRFAMS: TIGR00231; small_GTP; 1.

```

```

DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 395 AA; 43290 MW; 71B08165E7FB42C5 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 395;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
Db 3 KAKFERTKPHVNIQTIGHV 21

RESULT 13
EFTU_CORGL
ID EFTU_CORGL STANDARD; PRT; 396 AA.
AC P42439;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR CGL0497.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=94368062; PubMed=8085791;
RA Ludwig W., Neumaier J., Klugbauer N., Brockmann E., Rollier C.,
RA Klugbauer S., Reetz K., Schachtner I., Ludwigsen A.,
RA Bachleitner M., Fischer U., Schleifer K.H.;
RT "Phylogenetic relationships of Bacteria based on comparative sequence
RT analysis of elongation factor Tu and ATP-synthase beta-subunit
RL genes.";
RL Antonie Van Leeuwenhoek 64:285-305(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX Nakagawa S.;
RA "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X77034; CAA54323.1;
CC EMBL: AP005275; BAB97890.1;
CC HSP: P02990; IETU.
CC InterPro: IPR004541; EF-Tu.
CC InterPro: IPR004160; EFTU_Cterm.
CC InterPro: IPR004161; EFTU_D2.
CC InterPro: IPR00795; EF_GTPbind.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00009; GTP_EFTU; 1.

```

```

DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFACT.
DR TIGRFAMS; TIGR00231; small_gtp; 1.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
SQ SEQUENCE 396 AA; 43852 MW; F46333EC113AB5FD CRC64;

Query Match 76.7%; Score 69; DB 1; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RKGFKTPGVNIGTIXV 19
DB 3 KAKFKTPKPHVNIPTGHV 21

RESULT 14
EFTU_MYCLE
ID EFTU_MYCLE STANDARD; PRT; 396 AA.
AC P30768;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR MLI877.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A. PubMed=8446028;
RX MEDLINE=93341950; PubMed=8341612;
RA Silbak F., Bercovier H.;
RT "Nucleotide sequence of Mycobacterium leprae elongation factor
(EF-Tu) gene.";
RL Nucleic Acids Res. 21:3327-3327(1993).
RN [2]
RP SEQUENCE FROM N.A. PubMed=8446028;
RX MEDLINE=93188701; PubMed=8446028;
RA Honore N.T., Bergh S., Chanteau S., Doucet-Populaire F.,
RA Eigmeier K., Garnier T., Georges C., Launois P., Limpalaboon T.,
RA Newton S., Nsang K., del Portillo P., Ramesh G.R., Reddi P.,
RA Ridei P.R., Sittisombut N., Wu-Hunter S., Cole S.T.;
RT "Nucleotide sequence of the first cosmid from the Mycobacterium
leprae genome project: structure and function of the Rif-Str
regions.";
RL Mol. Microbiol. 7:207-214(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94375410; PubMed=8089081;
RA Dhandayuthapani S., Banu J.M., Kashiwabara Y.;
RT "Cloning and sequence determination of the gene coding for the
elongation factor Tu of Mycobacterium leprae.";
RL J. Biochem. 115:664-669(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";

```

```

RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L13276; AAA71969.1; -
CC EMBL; Z14314; CAA78674.1; -
CC EMBL; D13869; BAA02982.2; -
CC EMBL; AL583923; CAC30831.1; -
CC DR PIR; S34954; S34954.
CC DR HSP; P02990; LETU.
CC DR Leproma; MLI877; -
CC DR InterPro; IPR004541; EF-Tu.
CC DR InterPro; IPR004160; EFTU_Cterm.
CC DR InterPro; IPR004161; EFTU_D2.
CC DR InterPro; IPR007095; EF_GTPbind.
CC DR InterPro; IPR005225; Small_GTP.
CC DR Pfam; PF00009; GTP_EFTU; 1.
CC DR Pfam; PF03143; GTP_EFTU_D3; 1.
CC DR Pfam; PF03144; GTP_EFTU_D2; 1.
CC DR TIGRFAMS; TIGR00231; small_gtp; 1.
CC DR TIGRFAMS; TIGR00485; EF-Tu; 1.
CC DR PROSITE; PS00301; EFATOR_GTP; 1.
CC Complete proteome.
KW Elongation factor; Protein biosynthesis; GTP-binding;
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
FT CONFLICT 129 152 GVPYLVALKSDAVDDLELLV -> VYLTSWHLTSPT
FT CONFLICT 281 281 L -> V (IN REF. 2).
FT CONFLICT 349 349 G -> A (IN REF. 2).
FT CONFLICT 384 384 R -> P (IN REF. 3).
SQ SEQUENCE 396 AA; 43667 MW; D9CB88343C642778 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RKGFKTPGVNIGTIXV 19
DB 3 KAKFKTPKPHVNIPTGHV 21

RESULT 15
EFTU_STRCJ
ID EFTU_STRCJ STANDARD; PRT; 397 AA.
AC P95724;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Streptomyces cinnamonensis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=53446;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUE89;
RA Cappellano C., Monti F., Sosio M., Donadio S., Sarubbi E.;

```

Fri May 2 08:51:22 2003

Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
 AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
 BIOSYNTHESIS.
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 EF-TU/EF-1A SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; X98831; CAA67349.1; -
 CC HSSP; P02990; LETU.
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTUL Cterm.
 DR InterPro; IPR004161; EFTUL D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFACT.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW Elongation factor; Protein biosynthesis; GTP-binding.
 FT NP_BIND 19 26 GTP (BY SIMILARITY).
 FT NP_BIND 83 87 GTP (BY SIMILARITY).
 ET NP_BIND 138 141 GTP (BY SIMILARITY).
 SQ SEQUENCE 397 AA; 43891 MW; EEB8235730D8BBB7 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 397;
 Best Local Similarity 73.7%; Pred. No. 0.00012;
 . Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RQKFKTKPGVNIQTIXV 19
 : ||||| ||||| |
 Db 3 KAKFKTKPHVNIQTIGHV 21

Search completed: April 29, 2003, 09:40:54
 Job time : 26 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:36:44 ; Search time 29 Seconds
(without alignments)
134.996 Million cell updates/sec

Title: US-09-810-764A-5

Perfect score: 90

Sequence: 1 RGFERTKPGVNGITIXV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	86.7	467	10 Q9SEF8	Q9sef8 oryza sativ
2	78	86.7	467	10 Q8W2C3	Q8w2c3 oryza sativ
3	75	83.3	391	2 Q9F9S8	Q9f9s8 edta-degrad
4	72	80.0	409	16 Q8YP63	Q8yp63 anabaena sp
5	72	80.0	457	10 Q4O450	Q4o450 nicotiana s
6	72	80.0	474	10 Q9AXU2	Q9axu2 pelargonium
7	72	80.0	485	10 Q43364	Q43364 nicotiana s
8	69	76.7	391	16 Q8YHP2	Q8yhp2 bruceella me
9	69	76.7	394	2 Q933R9	Q933r9 ureaplasma
10	69	76.7	394	16 Q92GW4	Q92gw4 rickettsia
11	69	76.7	397	2 Q9RG55	Q9rg55 streptomyce
12	69	76.7	406	16 Q8YHQ4	Q8yhq4 bruceella me
13	68	75.6	397	2 Q9RG53	Q9rg53 streptover
14	68	75.6	397	2 Q93T39	Q93t39 streptomyce
15	67	74.4	397	16 Q97EH5	Q97eh5 clostridium
16	67	74.4	410	8 Q9TLV8	Q9tlv8 cyanidium c

17	66	73.3	394	16 Q9KV37	Q9kv37 vibrio chol
18	66	73.3	394	16 Q8KU26	Q8kuz6 vibrio chol
19	66	73.3	394	16 Q8ZJB2	Q8zjb2 versinia pe
20	66	73.3	394	16 Q8ZAN8	Q8zan8 versinia pe
21	66	73.3	394	16 Q8XED3	Q8xed3 escherichia
22	66	73.3	394	16 Q8X4S9	Q8x4s9 escherichia
23	66	73.3	396	16 Q8XGZ0	Q8xgz0 ralstonia s
24	66	73.3	400	16 Q8R7V2	Q8r7v2 thermoanaer
25	66	73.3	400	16 Q8R7T8	Q8r7t8 thermoanaer
26	66	73.3	409	8 Q9TJQ8	Q9tjq8 prototheca
27	65	72.2	397	16 Q8XFF8	Q8xfp8 clostridium
28	64	71.1	391	16 Q8UE16	Q8uel6 agrobacteri
29	64	71.1	391	16 Q925Y6	Q925y6 rhizobium m
30	63	70.0	393	2 Q9RAJ9	Q9raja moraxella s
31	62	68.9	394	16 Q9K117	Q9k1l7 neisseria m
32	62	68.9	394	16 Q9JR15	Q9jri5 neisseria m
33	62	68.9	396	16 Q99QM0	Q99qm0 caulobacter
34	62	68.9	440	3 Q8TGG6	Q8tgg6 aspergillus
35	60	66.7	395	16 Q8R603	Q8r6g3 fusobacteri
36	60	66.7	395	16 Q927I6	Q927i6 listeria in
37	60	66.7	395	16 Q8Y422	Q8y422 listeria mo
38	59	65.6	437	3 Q9C2P1	Q9c2p1 neurospora
39	57	63.3	398	16 Q97PV3	Q97pv3 streptococc
40	56	62.2	300	5 Q95TV3	Q95tv3 drosophila
41	56	62.2	401	2 Q9XD38	Q9xd38 leptospira
42	56	62.2	489	5 Q9V6Q7	Q9v6q7 drosophila
43	55	61.1	453	10 Q8W2C4	Q8w2c4 oryza sativ
44	53	58.9	358	2 Q9X6G1	Q9x6g1 coxiella bu
45	52	57.8	395	2 Q9R420	Q9r420 porphyromon

ALIGNMENTS

RESULT 1

Q9SEF8 PRELIMINARY; PRT; 467 AA.
ID AC Q9SEF8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast translational elongation factor Tu.
GN TUFA.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, JAPONICA NIPPONBARE;
RX MEDLINE-20062172; PubMed=10597036;
RA Lee J.H., Lee J.W., Chung Y.Y., Paek K.H., Shin J.S., Yun C.H.,
RA Kim J.K.;
RT "Cloning and characterization of the chloroplast elongation factor EF-
Tu cDNA of Oryza sativa L.";
RL Mol. Cells 9:484-490(1999).
DR EMBL; AF145053; AAF15312.1; -.
DR HSSP; P07157; LAIP.
DR InterPro; IPR004541; EF-Tu.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU_1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
SQ SEQUENCE 467 AA; 50354 MW; 79A119279A7F1847 CRC64;

Query Match 86.7%; Score 78; DB 10; Length 467;
Best Local Similarity 84.2%; Pred. No. 2.3e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKGFKRTKPGVNICTIXXV 19
| | | | | | | | | | | | | | | |
Db 61 RKGFKRTKPHVNIGTIGHV 79

RESULT 2
Q8W2C3 PRELIMINARY; PRT; 467 AA.
AC Q8W2C3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Translational elongation factor Tu.
GN TUFA.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
NCBI_TaxID=4530;
RN NCBI_TaxID=4530;
RP SEQUENCE FROM N.A.
RA Lee J.-H., Kim J.-K.;
RL "Chloroplast translational elongation factor Tu gene of rice.";
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF327413; AAL37431.1; -
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR000795; EF-GTPbind.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCr.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
KW Elongation factor.
SQ SEQUENCE 467 AA; 50413 MW; BA6D9FABDF02CCA CRC64;

Query Match 86.7%; Score 78; DB 10; Length 467;
Best Local Similarity 84.2%; Pred. No. 2.3e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKGFKRTKPGVNICTIXXV 19
| | | | | | | | | | | | | | | |
Db 61 RKGFKRTKPHVNIGTIGHV 79

RESULT 3
Q9F9S8 PRELIMINARY; PRT; 391 AA.
AC Q9F9S8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Eftu.
GN EFTU.
OS EDTA-degrading bacterium BNCI.
OC Bacteria; Proteobacteria; alpha subdivision.
NCBI_TaxID=85561;
RN NCBI_TaxID=85561;
RP SEQUENCE FROM N.A.
RC STRAIN=BNCI.
RX MEDLINE=21091958; PubMed=11157232;
RA Bohuslavsek J., Payne J.W., Liu Y., Bolton H. Jr., Xun L.;
RT "Cloning, Sequencing, and Characterization of a Gene Cluster Involved
in EDTA Degradation from the Bacterium BNCI.";
RI Appl. Environ. Microbiol. 67:688-695(2001).
RL Appl.

SEQUENCE FROM N.A.
STRAIN=BNCI;
MEDLINE=21091959; PubMed=11157233;
Liu Y., Louie T.M., Payne J., Bohuslavsek J., Bolton H. Jr., Xun L.;
Identification, purification, and characterization of iminodiacetate
oxidase from the EDTA-degrading bacterium BNCI.";
Appl. Environ. Microbiol. 67:696-701(2001).
EMBL: AF176664; AAG09263.1; -
HSSP: P07157; IAIP.
InterPro: IPR004541; EF-Tu.
InterPro: IPR004160; EFTU_Cterm.
InterPro: IPR004161; EFTU_D2.
InterPro: IPR000795; EF-GTPbind.
InterPro: IPR005225; Small_GTP.
Pfam: PF00009; GTP_EFTU; 1.
Pfam: PF03144; GTP_EFTU_D2; 1.
Pfam: PF03143; GTP_EFTU_D3; 1.
PRINTS: PR00315; ELONGATNFCr.
TIGRFAMS: TIGR00485; EF-Tu; 1.
PROSITE: PS00301; EFATOR_GTP; 1.
GTP-binding.
SEQUENCE 391 AA; 42794 MW; E6B41737CCDD77AA6 CRC64;

Query Match 83.3%; Score 75; DB 2; Length 391;
Best Local Similarity 78.9%; Pred. No. 6.5e-05;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKGFKRTKPGVNICTIXXV 19
| | | | | | | | | | | | | | | |
Db 3 RKGFKRTKPHVNIGTIGHV 21

RESULT 4
Q8YP63 PRELIMINARY; PRT; 409 AA.
ID Q8YP63;
AC Q8YP63;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Translation elongation factor EF-Tu.
GN TUFA OR ALL4337.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
NCBI_TaxID=103690;
RN NCBI_TaxID=103690;
RP SEQUENCE FROM N.A.
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
Yasuda M., Tabata S.;
Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
DNA Res. 8:205-213(2001).
EMBL: AP003596; BAB76036.1; -
InterPro: IPR004541; EF-Tu.
InterPro: IPR004160; EFTU_Cterm.
InterPro: IPR004161; EFTU_D2.
InterPro: IPR000795; EF-GTPbind.
InterPro: IPR005225; Small_GTP.
Pfam: PF00009; GTP_EFTU; 1.
Pfam: PF03144; GTP_EFTU_D2; 1.
Pfam: PF03143; GTP_EFTU_D3; 1.
PRINTS: PR00315; ELONGATNFCr.
TIGRFAMS: TIGR00485; EF-Tu; 1.
TIGRFAMS: TIGR00231; small_GTP; 1.
PROSITE: PS00301; EFATOR_GTP; 1.
Elongation factor; Complete proteome.
SEQUENCE 409 AA; 44811 MW; 084708FBAA8CD9D3 CRC64;

Query Match 80.0%; Score 72; DB 16; Length 409;

Best Local Similarity 78.9%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNIPTIXV 19
Db 3 RAKFERTKPHVNIPTIGHV 21

RESULT 5

Q04050 ID Q04050 PRELIMINARY; PRT; 457 AA.
AC Q04050
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast elongation factor TUA (EF-TUA) precursor (Fragment).
GN TUFA.
OS Nicotiana sylvestris (Wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4096;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF.
RX MEDLINE=93363910; PubMed=8358028;
RA Murayama Y., Matsubayashi T., Sugita M., Sugliura M.;
RT "Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco: the existence of two chloroplast elongation factor Tu species";
RL Plant Mol. Biol. 22:767-774(1993).
DR EMBL: D11375; BAA01974.1; -.
DR HSSP: P02990; LEFU.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU Cterm.
DR InterPro: IPR004161; EFTU D2.
DR InterPro: IPR00795; EF-GTPbind.
DR InterPro: IPR005225; Small-GTP.
DR Pfam: PF00009; GTP-EFTU; 1.
DR Pfam: PF03144; GTP-EFTU_D2; 1.
DR Pfam: PF03143; GTP-EFTU_D3; 1.
DR TIGRFAMs: TIGR00485; EF-Tu; 1.
DR TIGRFAMs: TIGR00231; small-GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
FT TRANSIT 1 70
FT CHAIN 71 >457
FT NON_TER 457 457
FT SEQUENCE 457 AA; 49732 MW; 8A30C50DC137F755 CRC64;

Query Match
Best Local Similarity 80.0%; Score 72; DB 10; Length 457;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNIPTIXV 19
Db 72 RAKFERTKPHVNIPTIGHV 90

RESULT 6

Q9AXU2 ID Q9AXU2 PRELIMINARY; PRT; 474 AA.
AC Q9AXU2
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast translational elongation factor Tu.
GN TUFA.
OS Pelargonium graveolens (rose geranium).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Geraniales; Geraniaceae; Pelargonium.

OX NCBI_TaxID=73200;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20550852; PubMed=11101151;
RA Kang C.J., Lee M.G., Cho Y.S., Lee J.W., Kyung Y.J., Shin J.S.,
RA Kim E.S., Kim J.K.;
RT "Characterization of geranium (Pelargonium graveolens) chloroplast EF-Tu cDNA";
RT Tu cDNA.;
RL Mol. Cells 10:579-583(2000).
DR EMBL: AF234537; AAK08141.1; -.
DR HSSP: P02990; LEFU.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU Cterm.
DR InterPro: IPR004161; EFTU D2.
DR InterPro: IPR00795; EF-GTPbind.
DR InterPro: IPR005225; Small-GTP.
DR Pfam: PF00009; GTP-EFTU; 1.
DR Pfam: PF03144; GTP-EFTU_D2; 1.
DR Pfam: PF03143; GTP-EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMs: TIGR00485; EF-Tu; 1.
DR TIGRFAMs: TIGR00231; small-GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
SQ SEQUENCE 474 AA; 51315 MW; A321742468553B65 CRC64;

Query Match
Best Local Similarity 80.0%; Score 72; DB 10; Length 474;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNIPTIXV 19
Db 68 RAKFERTKPHVNIPTIGHV 86

RESULT 7

Q43364 ID Q43364 PRELIMINARY; PRT; 485 AA.
AC Q43364
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast elongation factor TUB(EF-TUB) precursor.
GN TUFA OR TUBF.
OS Nicotiana sylvestris (Wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94373864; PubMed=8087886;
RA Sugita M., Murayama Y., Sugliura M.;
RT "Structure and differential expression of two distinct genes encoding the chloroplast elongation factor Tu in tobacco";
RL Curr. Genet. 25:164-168(1994).
RN [2]
RP SEQUENCE OF 61-485 FROM N.A.
RC TISSUE=YOUNG LEAF;
RX MEDLINE=93363910; PubMed=8358028;
RA Murayama Y., Matsubayashi T., Sugita M., Sugliura M.;
RT "Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco: the existence of two chloroplast elongation factor Tu species";
RL Plant Mol. Biol. 22:767-774(1993).
DR EMBL: D11470; BAA02028.1; -.
DR EMBL: D11376; BAA01975.1; -.
DR HSSP: P02990; LEFU.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU Cterm.
DR InterPro: IPR004161; EFTU D2.
DR InterPro: IPR00795; EF-GTPbind.
DR InterPro: IPR005225; Small-GTP.

Fri May 2 08:51:23 2003

us-09-810-764a-5.rspt

DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFACT.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 DR Elongation factor; GTP-binding; Protein biosynthesis; Transit peptide.
 KW POTENTIAL.
 FT TRANSIT 1 77
 FT CHAIN 78 485
 FT CHLOROPLAST ELONGATION FACTOR TUB
 FT (EF-TUB).
 SQ SEQUENCE 485 AA; 52688 MW; FE88EF5D48B92D29 CRC64;

Query Match 80.0%; Score 72; DB 10; Length 485;
 Best Local Similarity 78.9%; Pred. No. 0.00028;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKERTKPGVNIPTIXV 19
 ||||| || ||||| |
 Db 79 RGKFKERKPHVNIPTIGHV 97

RESULT 8

QYHP2 PRELIMINARY; PRT; 391 AA.
 AC Q8YHP2;
 DT 01-NAR-2002 (TREMBlrel. 20, Created)
 DT 01-WAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Protein translation elongation factor Tu (EF-TU).
 GN BME10755.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujar C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis";
 Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 RL EMBL; AE009518; AAL51936.1; -
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFACT.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 42604 MW; ED4CDF37183A900E CRC64;

Query Match 76.7%; Score 69; DB 16; Length 391;
 Best Local Similarity 73.7%; Pred. No. 0.00076;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKERTKPGVNIPTIXV 19
 : ||||| ||||| |
 Db 3 KSKFKERKPHVNIPTIGHV 21

RESULT 9

Q933R9 PRELIMINARY; PRT; 394 AA.
 AC Q933R9;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Elongation factor Tu.
 GN Ureaplasma urealyticum (Ureaplasma urealyticum biotype 2).
 OS Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=2130;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=VARIOUS STRAINS;
 RA Kong F., Gilbert G.L.;
 RT "Genomic based mapping and sequencing of U. parvum and U.
 urealyticum";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF270767; AAK58630.1; -
 DR EMBL; AF270758; AAK58621.1; -
 DR EMBL; AF270759; AAK58622.1; -
 DR EMBL; AF270760; AAK58623.1; -
 DR EMBL; AF270761; AAK58624.1; -
 DR EMBL; AF270762; AAK58625.1; -
 DR EMBL; AF270763; AAK58626.1; -
 DR EMBL; AF270764; AAK58627.1; -
 DR EMBL; AF270765; AAK58628.1; -
 DR EMBL; AF270766; AAK58629.1; -
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; UNKNOWN_1.
 KW GTP-binding.
 SQ SEQUENCE 394 AA; 42918 MW; 4CB4B2D776A5B145 CRC64;

Query Match 76.7%; Score 69; DB 2; Length 394;
 Best Local Similarity 73.7%; Pred. No. 0.00077;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKERTKPGVNIPTIXV 19
 : ||||| ||||| |
 Db 3 KAKFKERKPHVNIPTIGHV 21

RESULT 10

Q92GW4 PRELIMINARY; PRT; 394 AA.
 AC Q92GW4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Elongation factor EF-Tu.
 GN TUF OR RC1008.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsia.
 OX NCBI_TaxID=781;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=WALISH 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";

Science 293:2093-2098(2001).
 EMBL; AE008652; AAL03546.1; -
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.
 InterPro; IPR000795; EF_GTPbind.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00009; GTP_EFTU; 1.
 Pfam; PF03144; GTP_EFTU_D2; 1.
 Pfam; PF03143; GTP_EFTU_D3; 1.
 TIGRFAMS; TIGR00485; EF-Tu; 1.
 TIGRFAMS; TIGR00231; small_GTP; 1.
 PROSITE; PS00301; EFATOR_GTP; UNKNOWN_1.
 Complete proteome.
 NCBI_TaxID=29459;
 SEQUENCE FROM N.A.
 STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 MEDLINE=20020109; PubMed=11756688;
 DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mijer C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 Haselkorn R., Kyrpides N., Overbeek R.;
 "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis.";
 Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 EMBL; AE009516; AAL51923.1; -
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.
 InterPro; IPR000795; EF_GTPbind.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00009; GTP_EFTU; 1.
 Pfam; PF03144; GTP_EFTU_D2; 1.
 Pfam; PF03143; GTP_EFTU_D3; 1.
 PRINTS; PR00315; ELONGATNFT.
 TIGRFAMS; TIGR00485; EF-Tu; 1.
 TIGRFAMS; TIGR00231; small_GTP; 1.
 PROSITE; PS00301; EFATOR_GTP; 1.
 Complete proteome.
 NCBI_TaxID=55404;
 SEQUENCE FROM N.A.
 STRAIN=TU1063;
 Olshoorn-Tielemans L.N., Claij N., Hilgenfeld R., Kraal B.;
 "Elongation factor Tu from the pulvomycin producer Streptomyces
 mobaraense is resistant to both pulvomycin and kirromycin.";
 Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 EMBL; AF153618; AAF22608.1; -
 HSSP; P02990; 1ETU.
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.

Query Match 76.7%; Score 69; DB 2; Length 397;
 Best Local Similarity 73.7%; Pred. No. 0.00078;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RGFERTKPGVNIPTIXV 19
 : ||||| |||||
 Db 3 KAKFERTKPHVNIPTIGHV 21

RESULT 11

Q9RG55 PRELIMINARY; PRT; 397 AA.

AC Q9RG55;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Elongation factor Tu.
 GN Tuf1.

OS Streptomyces mobaraensis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=35621;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC29032;
 RA Olshoorn-Tielemans L.N., Claij N., Hilgenfeld R., Kraal B.;
 "Elongation factor Tu from the pulvomycin producer Streptomyces
 mobaraense is resistant to both pulvomycin and kirromycin.";
 Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF153617; AAF22606.1; -
 DR HSSP; P02990; 1ETU.
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFT.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW GTP-binding.
 SQ SEQUENCE 397 AA; 43811 MW; DA439151BFE6BC2E CRC64;

Query Match 76.7%; Score 69; DB 2; Length 397;
 Best Local Similarity 73.7%; Pred. No. 0.00078;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RGFERTKPGVNIPTIXV 19
 : ||||| |||||
 Db 3 KAKFERTKPHVNIPTIGHV 21

RESULT 12

Q8YHQ4

ID Q8YHQ4 PRELIMINARY; PRT; 406 AA.
 AC Q8YHQ4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Protein translation elongation factor Tu (EF-Tu).
 GN EME10742.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mijer C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 Haselkorn R., Kyrpides N., Overbeek R.;
 "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis.";
 Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 EMBL; AE009516; AAL51923.1; -
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.
 InterPro; IPR000795; EF_GTPbind.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00009; GTP_EFTU; 1.
 Pfam; PF03144; GTP_EFTU_D2; 1.
 Pfam; PF03143; GTP_EFTU_D3; 1.
 PRINTS; PR00315; ELONGATNFT.
 TIGRFAMS; TIGR00485; EF-Tu; 1.
 TIGRFAMS; TIGR00231; small_GTP; 1.
 PROSITE; PS00301; EFATOR_GTP; 1.
 Complete proteome.
 NCBI_TaxID=55404;
 SEQUENCE FROM N.A.
 STRAIN=TU1063;
 Olshoorn-Tielemans L.N., Claij N., Hilgenfeld R., Kraal B.;
 "Elongation factor Tu from the pulvomycin producer Streptomyces
 mobaraense is resistant to both pulvomycin and kirromycin.";
 Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 EMBL; AF153618; AAF22608.1; -
 HSSP; P02990; 1ETU.
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.

Query Match 76.7%; Score 69; DB 16; Length 406;
 Best Local Similarity 73.7%; Pred. No. 0.0008;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RGFERTKPGVNIPTIXV 19
 : ||||| |||||
 Db 18 KSKFERTKPHVNIPTIGHV 36

RESULT 13

Q9RG53 PRELIMINARY; PRT; 397 AA.

AC Q9RG53;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Elongation factor Tu.
 GN Tuf1.

OS Streptococcus thermophilus netropsis (Streptococcus thermophilus).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=55404;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TU1063;
 Olshoorn-Tielemans L.N., Claij N., Hilgenfeld R., Kraal B.;
 "Elongation factor Tu from the pulvomycin producer Streptomyces
 mobaraense is resistant to both pulvomycin and kirromycin.";
 Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 EMBL; AF153618; AAF22608.1; -
 HSSP; P02990; 1ETU.
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.

```
DR InterPro: IPR000795; EF-GTPbind.
DR Pfam: PF00009; GTP_EFTU_D2; 1.
DR Pfam: PF03144; GTP_EFTU_D3; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR GTP-binding.
KW GTP-binding.
SQ SEQUENCE 397 AA; 43682 MW; 5F3A81F3FC42914F CRC64;

Query Match 75.6%; Score 68; DB 2; Length 397;
Best Local Similarity 68.4%; Pred. No. 0.0012;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 KAKFERTKPHVNIQTIGHI 21

RESULT 14
Q3T39 PRELIMINARY; PRT; 397 AA.
AC Q3T39;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Elongation factor Tu.
GN TUFI.
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=84/25;
RA Kormanec J., Nguyen L.D., Novotna J., Knirschova R., Weiser J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF368284; AAK54131.1; -.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
KW GTP-binding.
SQ SEQUENCE 397 AA; 43683 MW; BF107EA8036FEF82 CRC64;

Query Match 75.6%; Score 68; DB 2; Length 397;
Best Local Similarity 68.4%; Pred. No. 0.0012;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 KAKFERTKPHVNIQTIGHI 21

RESULT 15
Q97EH5 PRELIMINARY; PRT; 397 AA.
AC Q97EH5;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Elongation factor Tu (Ef-Tu).
GN CAC3136.
OS Clostridium acetobutylicum.

Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
Clostridiales; Clostridiaceae; Clostridium.
NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Oiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabath F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RT J. Bacteriol. 183:4823-4838(2001).
RL EMBL: AE007809; AAK81075.1; -.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
KW Complete proteome.
SQ SEQUENCE 397 AA; 43452 MW; F031B2F700434452 CRC64;

Query Match 74.4%; Score 67; DB 16; Length 397;
Best Local Similarity 82.4%; Pred. No. 0.0018;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXV 19
DB 5 KFERTKPHVNIQTIGHV 21

Search completed: April 29, 2003, 09:41:32
Job time : 32 secs
```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:39:04 ; Search time 29 Seconds
(without alignments)
19.277 Million cell updates/sec

Title: US-09-810-764a-5

Perfect score: 90

Sequence: 1 RGFERTKPGVNIQTIXV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	75.6	409	4	US-09-140-466-3
2	67	74.4	408	4	US-09-140-466-4
3	66	73.3	394	4	US-09-140-466-5
4	55	61.1	409	4	US-09-140-466-6
5	48	53.3	394	4	US-09-140-466-7
6	48	53.3	414	4	US-09-134-001C-5377
7	41	45.6	861	1	US-08-764-100-24
8	40	44.4	406	1	US-07-973-431B-1
9	39	43.3	260	4	US-09-254-465A-23
10	39	43.3	263	4	US-09-254-465A-25
11	39	43.3	299	4	US-09-188-930-189
12	39	43.3	299	4	US-09-188-930-331
13	39	43.3	299	4	US-09-462-270-2
14	39	43.3	299	4	US-09-254-465A-1
15	39	43.3	739	4	US-08-444-818-148
16	39	43.3	2955	2	US-08-443-260-3
17	39	43.3	2955	3	US-08-442-805A-3
18	39	43.3	2955	3	US-08-443-900A-3
19	39	43.3	2955	4	US-08-444-818-124
20	39	43.3	2955	4	US-08-249-843-3
21	39	43.3	3011	2	US-08-444-818-138
22	39	43.3	3011	4	US-08-833-678A-6
23	39	43.3	3011	4	US-08-444-818-177
24	39	43.3	3011	4	US-08-529-169A-6
25	38	42.2	255	1	US-08-152-019A-36
26	38	42.2	410	4	US-09-140-466-2
27	38	42.2	461	4	US-09-134-001C-3316

```

28 38 42.2 1112 2 US-08-714-402-2 Sequence 2, Appli
29 38 42.2 1161 4 US-09-327-536-2 Sequence 2, Appli
30 37 41.1 330 2 US-08-997-080-30 Sequence 30, Appl
31 37 41.1 330 2 US-08-997-362-30 Sequence 30, Appl
32 37 41.1 330 3 US-08-873-970-30 Sequence 30, Appl
33 37 41.1 330 4 US-09-095-855-30 Sequence 30, Appl
34 37 41.1 330 4 US-08-705-347A-30 Sequence 30, Appl
35 37 41.1 330 4 US-09-324-542-30 Sequence 30, Appl
36 37 41.1 330 4 US-09-205-426-30 Sequence 30, Appl
37 37 41.1 330 4 US-09-200-643-30 Sequence 30, Appl
38 37 41.1 1239 2 US-08-937-931-2 Sequence 2, Appli
39 37 41.1 1239 4 US-09-285-502-2 Sequence 2, Appli
40 37 41.1 1239 4 US-09-709-126-2 Sequence 2, Appli
41 37 41.1 1239 4 US-09-871-385A-2 Sequence 2, Appli
42 36.5 40.6 765 4 US-08-444-818-70 Sequence 70, Appl
43 36 40.0 28 1 US-07-696-551B-1 Sequence 1, Appli
44 36 40.0 315 4 US-09-154-750A-91 Sequence 91, Appl
45 35 38.9 185 2 US-08-770-544-18 Sequence 18, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-140-466-3
; Sequence 3, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW PJC DP
; CURRENT APPLICATION NUMBER: US/09/140,466
; EARLIER FILING DATE: 1998-08-26
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Anacystis nidulans
US-09-140-466-3

```

```

Query Match 75.6%; Score 68; DB 4; Length 409;
Best Local Similarity 73.7%; Pred. No. 0.0012;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 RAKFERTKPHANIGTIGHV 21

```

```

RESULT 2
US-09-140-466-4
; Sequence 4, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW PJC DP
; CURRENT APPLICATION NUMBER: US/09/140,466
; EARLIER FILING DATE: 1998-08-26
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14

```

US-09-140-466-5

FACEIT NO. 5773700
: GENERAL INFORMATION:

: GENERAL INFORMATION:
 : APPLICANT: Lu, Yinchun
 : APPLICANT: Haseltine, William A
 : TITLE OF INVENTION: YC1 Protein, Gene, And Uses Thereof
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: David G. Conlin; Dike, Bronstein,
 : ADDRESSEE: Roberts & Cushman
 : STREET: 130 Water Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02109

1 RGKFERTKPGVN 12

Query Match 43.3%; Score 39; DB 4; Length 299;
 Best Local Similarity 58.3%; Pred. No. 54;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVN 12
 |||:|||||:
 Db 263 RGHFDRTKKGS 274

RESULT 12
 US-09-188-930-331
 ; Sequence 331, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 331
 ; LENGTH: 299
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-188-930-331

Query Match 43.3%; Score 39; DB 4; Length 299;
 Best Local Similarity 58.3%; Pred. No. 54;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVN 12
 |||:|||||:
 Db 263 RGHFDRTKKGS 274

RESULT 13
 US-09-462-270-2
 ; Sequence 2, Application US/09462270
 ; Patent No. 6358707
 ; GENERAL INFORMATION:
 ; APPLICANT: Smithkline Beecham Corporation
 ; TITLE OF INVENTION: Human FII Antigen: A Cell Surface
 ; TITLE OF INVENTION: Receptor Involved in Platelet Aggregation
 ; FILE REFERENCE: GH-70150US
 ; CURRENT APPLICATION NUMBER: US/09/462,270
 ; CURRENT FILING DATE: 2000-01-05
 ; PRIOR APPLICATION NUMBER: 60/052,186
 ; PRIOR FILING DATE: 1997-07-10
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 299
 ; TYPE: PRT
 ; ORGANISM: HOMO SAPIENS
 ; US-09-462-270-2

Query Match 43.3%; Score 39; DB 4; Length 299;
 Best Local Similarity 58.3%; Pred. No. 54;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVN 12
 |||:|||||:
 Db 263 RGHFDRTKKGS 274

RESULT 14
 US-09-254-465A-1

Db 240 RGHFDRTKKGS 251

RESULT 10
 US-09-254-465A-25
 ; Sequence 25, Application US/09254465A
 ; Patent No. 6410708
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: COMPOUNDS, COMPOSITIONS AND METHODS FOR THE TREATMENT
 ; TITLE OF INVENTION: OF DISEASES CHARACTERIZED BY A33- RELATED ANTIGENS
 ; FILE REFERENCE: P12181(US)
 ; CURRENT APPLICATION NUMBER: US/09/254,465A
 ; CURRENT FILING DATE: 1999-03-05
 ; PRIOR APPLICATION NUMBER: PCT/US98/24855
 ; PRIOR FILING DATE: 1998-11-20
 ; PRIOR APPLICATION NUMBER: US 60/066,364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: US 60/078,936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: PCT/US98/19437
 ; PRIOR FILING DATE: 1998-09-17
 ; * NUMBER OF SEQ ID NOS: 30
 ; SEQ ID NO 25
 ; LENGTH: 263
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-254-465A-25

Query Match 43.3%; Score 39; DB 4; Length 263;
 Best Local Similarity 58.3%; Pred. No. 47;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVN 12
 |||:|||||:
 Db 243 RGHFDRTKKGS 254

RESULT 11
 US-09-188-930-189
 ; Sequence 189, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 189
 ; LENGTH: 299
 ; TYPE: PRT
 ; ORGANISM: Human

Query Match 43.3%; Score 39; DB 4; Length 299;
 Best Local Similarity 58.3%; Pred. No. 54;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

NAME/KEY: UNSURE
 LOCATION: (247)...(247)
 NAME/KEY: UNSURE
 LOCATION: (289)...(289)
 US-09-188-930-189

; Sequence 1, Application US/09254465A
; Patent No. 6410708
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Fong, Sherman
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Napier, Mary A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOUNDS, COMPOSITIONS AND METHODS FOR THE TREATMENT
; FILE REFERENCE: P1216R1(US)
; CURRENT APPLICATION NUMBER: US/09/254,465A
; PRIOR FILING DATE: 1999-03-05
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US98/24855
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: US 60/066,364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 60/078,936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 30
; SEQ ID NO 1
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-465A-1

Query Match 43.3%; Score 39; DB 4; Length 299;
Best Local Similarity 58.3%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFERTKPGVN 12
|||:|||||
Db 263 RGHFDRKKGTS 274

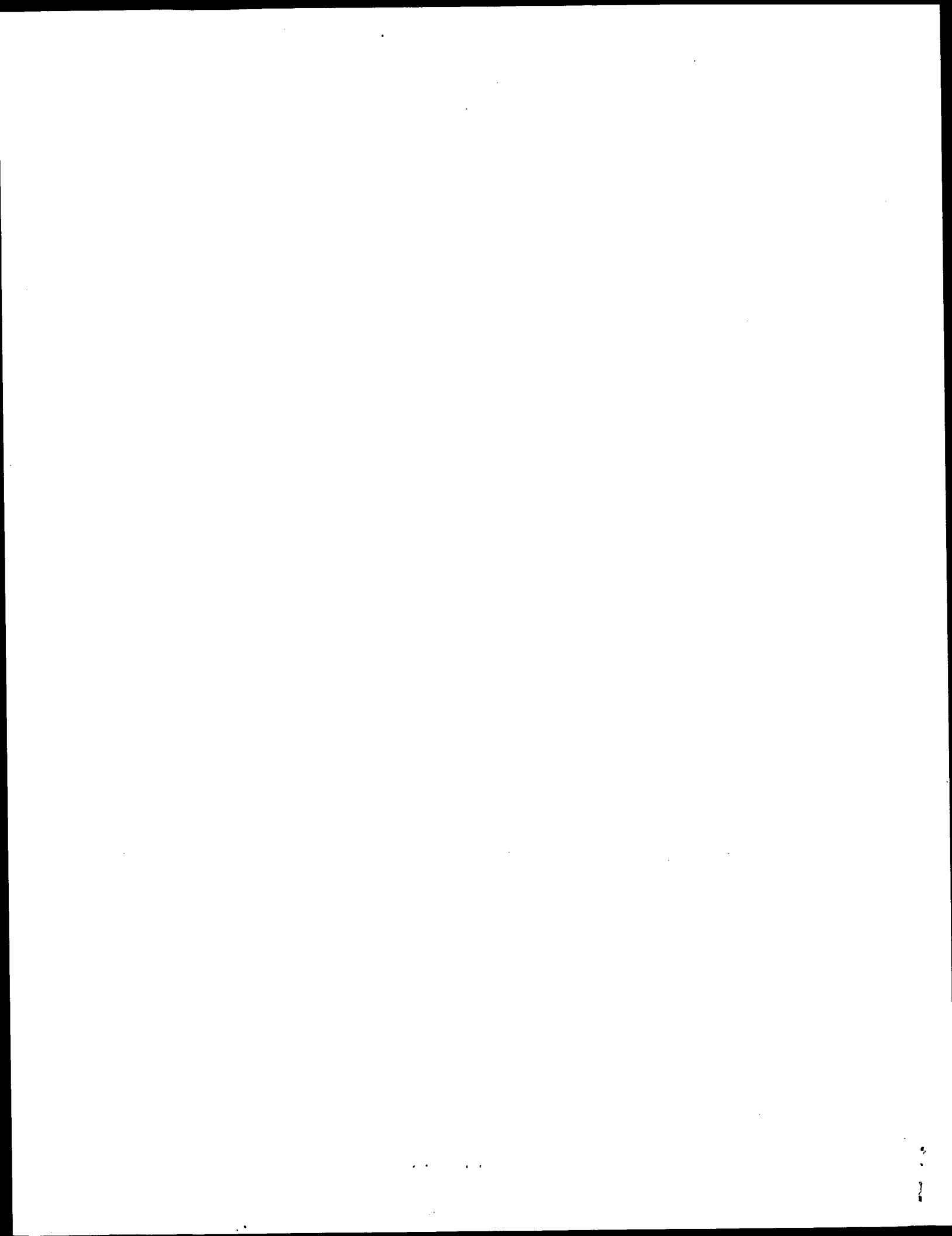
RESULT 15
US-08-444-818-148
; Sequence 148, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-444-818-148

Query Match 43.3%; Score 39; DB 4; Length 739;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFERTKPGVN 12
|||:|||||
Db 297 RGRGTGRKPGIN 308

Search completed: April 29, 2003, 09:42:57
Job time : 31 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:41:35 ; Search time 15 Seconds
(without alignments)
101.498 Million cell updates/sec

Title: US-09-810-764A-5
Perfect score: 90
Sequence: 1 RGKFKTPGVNIGTIXXV 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	76.7	395	9	US-09-738-626-6937
2	68	75.6	409	10	US-09-845-335-3
3	67	74.4	408	10	US-09-845-335-4
4	66	73.3	394	10	US-09-912-020-247
5	66	73.3	394	10	US-09-845-335-6
6	66	73.3	394	10	US-09-815-242-10362
7	66	73.3	394	10	US-09-815-242-10431
8	66	73.3	394	10	US-09-815-242-11059
9	66	73.3	394	10	US-09-815-242-11069
10	66	73.3	394	10	US-09-815-242-13964
11	62	68.9	22	8	US-08-831-310-9
12	62	68.9	22	10	US-09-488-737-1
13	62	68.9	399	8	US-08-831-310-4
14	62	68.9	399	10	US-09-815-242-11415
15	61	67.8	397	10	US-09-815-242-11995
16	61	67.8	397	10	US-09-815-242-12002
17	57	63.3	398	10	US-09-815-242-13171
18	55	61.1	409	10	US-09-845-335-5
19	53	58.9	394	10	US-09-841-132-491

20	53	58.9	394	10	US-09-841-132-561	Sequence 561, Ap
21	53	58.9	394	12	US-10-007-693-68	Sequence 68, Appl
22	53	58.9	394	12	US-10-007-693-103	Sequence 103, Ap
23	52	57.8	395	10	US-09-815-242-4887	Sequence 4887, Ap
24	52	57.8	395	10	US-09-815-242-10482	Sequence 10482, A
25	48	53.3	394	10	US-09-815-242-5574	Sequence 5574, Ap
26	48	53.3	394	10	US-09-815-242-12442	Sequence 12442, A
27	45	50.0	479	10	US-09-925-300-1515	Sequence 1515, Ap
28	42	46.7	20	9	US-09-736-457-1847	Sequence 1847, Ap
29	42	46.7	20	9	US-09-902-941-1847	Sequence 1847, Ap
30	42	46.7	20	9	US-09-849-626-1847	Sequence 1847, Ap
31	42	46.7	20	9	US-10-017-754-1847	Sequence 1847, Ap
32	42	46.7	200	9	US-09-736-457-789	Sequence 324, Ap
33	42	46.7	200	9	US-09-736-457-789	Sequence 324, Ap
34	42	46.7	200	9	US-09-902-941-324	Sequence 324, Ap
35	42	46.7	200	9	US-09-902-941-789	Sequence 789, Ap
36	42	46.7	200	9	US-09-849-626-324	Sequence 324, Ap
37	42	46.7	200	9	US-09-849-626-789	Sequence 789, Ap
38	42	46.7	200	9	US-10-017-754-324	Sequence 324, Ap
39	42	46.7	200	9	US-10-017-754-789	Sequence 789, Ap
40	42	46.7	201	10	US-09-214-881A-10	Sequence 10, Appl
41	42	46.7	207	9	US-09-736-457-1667	Sequence 1667, Ap
42	42	46.7	207	9	US-09-902-941-1667	Sequence 1667, Ap
43	42	46.7	207	9	US-09-902-941-1913	Sequence 1913, Ap
44	42	46.7	207	9	US-09-849-626-1667	Sequence 1667, Ap
45	42	46.7	207	9	US-09-849-626-1913	Sequence 1913, Ap

ALIGNMENTS

RESULT 1
US-09-738-626-6937
; Sequence 6937, Application US/09738626
; Publication No. US20020197605A1

; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOHI, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6937
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6937

Query Match 76.7%; Score 69; DB 9; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.00054;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXXV 19

Db 3 KAKFKTKPVNIGTICHV 21

RESULT 2

US-09-845-335-3
 ; Sequence 3, Application US/09845335
 ; Patent No. US20020058266A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CLOUGH, BARBARA
 ; APPLICANT: PREISER, PETER
 ; APPLICANT: WILSON, ROBERT
 ; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
 ; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
 ; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
 ; FILE REFERENCE: N68837B GCW PJC DP
 ; CURRENT APPLICATION NUMBER: US/09/845,335
 ; CURRENT FILING DATE: 2001-12-26
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 409
 ; TYPE: PRT
 ; ORGANISM: Anacystis nidulans
 ; US-09-845-335-3

Query Match 75.6%; Score 68; DB 10; Length 409;
 Best Local Similarity 73.7%; Pred. No. 0.00083;
 Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy. 1 RGKFKRTKPGVNIQTIXXV 19
 | ||||| |||||
 Db 3 RAKFKTPHANIQTIGHV 21

RESULT 3

US-09-845-335-4
 ; Sequence 4, Application US/09845335
 ; Patent No. US20020058266A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CLOUGH, BARBARA
 ; APPLICANT: PREISER, PETER
 ; APPLICANT: WILSON, ROBERT
 ; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
 ; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
 ; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
 ; FILE REFERENCE: N68837B GCW PJC DP
 ; CURRENT APPLICATION NUMBER: US/09/845,335
 ; CURRENT FILING DATE: 2001-12-26
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 408
 ; TYPE: PRT
 ; ORGANISM: Cryptomonas phi
 ; US-09-845-335-4

Query Match 74.4%; Score 67; DB 10; Length 408;
 Best Local Similarity 73.7%; Pred. No. 0.0012;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RGKFKRTKPGVNIQTIXXV 19
 | ||||| |||||
 Db 3 RDKFKSRPHVNIQTIGHV 21

RESULT 4

US-09-912-020-247
 ; Sequence 247, Application US/09912020
 ; Patent No. US2002004592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Trawick, John
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Froelich, Jamie M.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
 ; TITLE OF INVENTION: ESCHERICHIA COLI
 ; FILE REFERENCE: ELITRA.001DVI
 ; CURRENT APPLICATION NUMBER: US/09/912,020
 ; CURRENT FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: 09/492,709
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: 60/117,405
 ; PRIOR FILING DATE: 1999-01-27
 ; NUMBER OF SEQ ID NOS: 485
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 247
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: E. Coli
 ; US-09-912-020-247

Query Match 73.3%; Score 66; DB 10; Length 394;
 Best Local Similarity 76.5%; Pred. No. 0.0018;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KFERTKPGVNIQTIXXV 19
 | ||||| |||||
 Db 5 KFERTKPHVNIQTIGHV 21

RESULT 5

US-09-845-335-6
 ; Sequence 6, Application US/09845335
 ; Patent No. US20020058266A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CLOUGH, BARBARA
 ; APPLICANT: PREISER, PETER
 ; APPLICANT: WILSON, ROBERT
 ; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
 ; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
 ; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
 ; FILE REFERENCE: N68837B GCW PJC DP
 ; CURRENT APPLICATION NUMBER: US/09/845,335
 ; CURRENT FILING DATE: 2001-12-26
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; US-09-845-335-6

Query Match 73.3%; Score 66; DB 10; Length 394;
 Best Local Similarity 76.5%; Pred. No. 0.0018;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KFERTKPGVNIQTIXXV 19
 | ||||| |||||
 Db 5 KFERTKPHVNIQTIGHV 21

RESULT 6

US-09-815-242-10362
 ; Sequence 10362, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10362
LENGTH: 394
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10362

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNGTIXXV 19
||||||| ||:||||
DB 5 KFERTKPHVNVGTIGHV 21

RESULT 7
US-09-815-242-10431
Sequence 10431, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10362
LENGTH: 394
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10362

NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10431
LENGTH: 394
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10431

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNGTIXXV 19
||||||| ||:||||
DB 5 KFERTKPHVNVGTIGHV 21

RESULT 8
US-09-815-242-11059
Sequence 11059, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11059
LENGTH: 394
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11059

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNGTIXXV 19
||||||| ||:||||
DB 5 KFERTKPHVNVGTIGHV 21

RESULT 9
US-09-815-242-11069
Sequence 11069, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIORITY FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11069
LENGTH: 394
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11069

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIPTIXV 19
||| ||| ||| ||| |||
Db 5 KFERTKPHVNVGTIGHV 21

RESULT 10
US-09-815-242-13964
Sequence 13964, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIORITY FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13964
LENGTH: 409
TYPE: PRT
ORGANISM: Salmonella typhi
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(409)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13964

Query Match 73.3%; Score 66; DB 10; Length 409;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIPTIXV 19
||| ||| ||| ||| |||
Db 20 KFERTKPHVNVGTIGHV 36

RESULT 11
US-08-831-310-9
Sequence 9, Application US/08831310
Patent No. US20020026035A1
GENERAL INFORMATION:
APPLICANT: Kleanthous, Harold et al.
TITLE OF INVENTION: Helicobacter GHPO 1360 and
TITLE OF INVENTION: GHPO 750 Polypeptides and Corresponding Polynucleotides
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/831,310
FILING DATE: 01-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,175
REFERENCE/DOCKET NUMBER: 06132/037001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-831-310-9

Query Match 68.9%; Score 62; DB 8; Length 22;
Best Local Similarity 76.5%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIPTIXV 19
||| ||| ||| ||| |||
Db 4 KFERTKPHVNVGTIGHV 20

```
RESULT 12
US-09-488-737-1
; Sequence 1, Application US/09488737
; Patent No. US20020151462A1
; GENERAL INFORMATION:
; APPLICANT: Lissolo, Ling
; TITLE OF INVENTION: Helicobacter Pylori Membrane Proteins
; FILE REFERENCE: 50019/005002
; CURRENT APPLICATION NUMBER: US/09/488,737
; CURRENT FILING DATE: 2000-01-20
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/FR96/01552
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: FR 95/11,890
; PRIOR FILING DATE: 1995-10-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-488-737-1

Query Match      68.9%; Score 62; DB 10; Length 22;
Best Local Similarity 76.5%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  3 KFERTKPGVNIQTIXV 19
Db  4 KFNRTKPHVNIQTIGHV 20

RESULT 13
US-08-831-310-4
; Sequence 4, Application US/08831310
; Patent No. US20020026035A1
; GENERAL INFORMATION:
; APPLICANT: Kleinhous, Harold et al.
; TITLE OF INVENTION: Helicobacter GHPO 1360 and
; TITLE OF INVENTION: GHPO 750 Polypeptides and Corresponding Polynucleotides
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,310
; FILING DATE: 01-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,175
; REFERENCE/DOCKET NUMBER: 06132/037001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NOS: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; FRAGMENT TYPE: Internal
US-08-831-310-4

Query Match      68.9%; Score 62; DB 8; Length 399;
Best Local Similarity 76.5%; Pred. No. 0.0086;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  3 KFERTKPGVNIQTIXV 19
Db  5 KFNRTKPHVNIQTIGHV 21

RESULT 14
US-09-815-242-11415
; Sequence 11415, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: EUTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11415
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11415

Query Match      68.9%; Score 62; DB 10; Length 399;
Best Local Similarity 76.5%; Pred. No. 0.0086;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  3 KFERTKPGVNIQTIXV 19
Db  5 KFNRTKPHVNIQTIGHV 21

RESULT 15
US-09-815-242-11995
; Sequence 11995, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
```

```

; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11995
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11995

```

```

Query Match      67.8%; Score 61; DB 10; Length 397;
Best Local Similarity 70.6%; Pred. No. 0.013;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 3 KEERTKPGVNIQTXXV 19
DQ 5 KPERNKPVNVGTIGHV 21

```

Search completed: April 29, 2003, 09:49:33
Job time : 15 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 1, 2003, 05:55:55 ; Search time 5099 Seconds
(without alignments)
9594.394 Million cell updates/sec

Title: US-09-810-764A-6
Perfect score: 1681
Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.hig.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pln.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1215.4	72.3	138931	2	AP004023	AP004023 Oryza sat
2	1210.6	72.0	2050	8	AF327413	AF327413 Oryza sat
3	1200.6	71.4	1678	8	AF145053	AF145053 Oryza sat
4	790.6	47.0	3079	8	GMRNAFTU	X89058 G.max DNA f
5	787.2	46.8	2313	8	GMRNAFTU	X69062 G.max tufa
6	774	46.0	1584	8	AF234537	AF234537 Pelargon
7	711.8	42.3	1731	8	PSY14561	Y14561 Pisum sativ
8	703.4	41.8	2342	8	TOBTFTU	M94304 Nicotiana t
9	700.2	41.7	2667	8	TOBTUFA2	D11469 Nicotiana s
10	687.2	40.9	2272	8	ATTUFA	X52256 A.thaliana
11	685.6	40.8	1663	8	AF419609	AF419609 Arabidops
12	685.6	40.8	1669	8	AF410329	AF410329 Arabidops
13	685.6	40.8	1688	8	AY074355	AY074355 Arabidops
14	685.6	40.8	10936	8	ATCF9F13	AL080253 Arabidops
15	685.6	40.8	198427	8	ATCHRIV52	AL161552 Arabidops
16	675.2	40.2	1436	8	TOBTUFB	D11376 N. sylvestr
17	675.2	40.2	3675	8	TOBTUFB2	D11470 Nicotiana s
18	658.4	39.2	1390	8	TOBTUFA	D11375 N. sylvestr
19	630.2	37.5	298750	1	AP005375	AP005375 Thermosyn
20	629.8	37.5	1337	8	GMTUFB2	Y15108 Glycine max
21	622.2	37.0	5139	1	ANORF150	X17442 A.nidulans
22	604.4	36.0	10492	1	AE001892	AE001892 Deinococc
23	604.2	35.9	9997	1	AE002041	AE002041 Deinococc
24	590.4	35.1	348550	1	AP003596	AP003596 Nostoc sp
25	565	33.6	143308	1	D90913	D90913 Synecocyst
26	557.6	33.2	1741	1	AF007125	AF007125 Streptomy
27	554.4	33.0	5273	1	SPSTR	X15646 Spirulina p
28	553.4	32.9	200799	8	AF137379	AF137379 Nephrosel
29	548.4	32.6	1358	1	S79408	S79408 tufl-elonga
30	548.4	32.6	23087	1	SCD40A	AL161691 Streptomy
31	547	32.5	1900	1	SCTUFLFUS	X77039 S.coelicolo
32	539.8	32.1	2012	1	SATUFBRNA	X82820 S.aurantia
33	538	32.0	1191	1	AB073986	AB073986 Myxococcu
34	538	32.0	1826	1	PRU67308	U67308 Planobispor
35	536.4	31.9	1191	1	TCHOESTUF	X76871 T.cuprinus
36	536.4	31.9	2742	1	PRFUSTUF	X98830 P.rosea fus
37	535.8	31.9	2250	1	AF153618	AF153618 Streptomy
38	533.6	31.7	1374	1	AF368284	AF368284 Streptomy
39	532.6	31.7	9150	1	AE011727	AE011727 Xanthomon
40	531.8	31.6	5203	1	TCU78300	U78300 Thiobacillu
41	531.8	31.6	13867	1	AE011726	AE011726 Xanthomon
42	529.8	31.5	13486	1	AE004842	AE004842 Pseudomon
43	529.4	31.5	2836	1	SRTUFL	X57057 S.ramocissi
44	526.4	31.3	11171	1	AE004843	AE004843 Pseudomon
45	524.8	31.2	7934	1	AE005800	AE005800 Caulobact

ALIGNMENTS

RESULT 1	AP004023	138931 bp	DNA	linear	HTG 21-MAR-2002
LOCUS	Oryza sativa (japonica cultivar-group)	chromosome 2 clone			
DEFINITION	OJ1126_D09, *** SEQUENCING IN PROGRESS ***				
ACCESSION	AP004023				
VERSION	AP004023.1	GI:15130685			
KEYWORDS	HTG; HTGS, PHASE2.				
SOURCE	Oryza sativa (japonica cultivar-group)	(cultivar:Nipponbare) DNA,			
ORGANISM	Clone:OJ1126_D09.				
	Oryza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Ehrhartoideae; Oryzeae; Oryza.				
REFERENCE	1				

AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC clone: OJ1126_D09
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 138931)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
COMMENT (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/) (Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
FEATURES Location/Qualifiers
 source 1..138931
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1126_D09"
BASE COUNT 39463 a 30336 c 30692 g 37874 t 566 others
ORIGIN
 Query Match 72.3%; Score 1215.4; DB 2; Length 138931;
 Best Local Similarity 84.4%; Pred. No. 3e-154;
 Matches 1405; Conservative 0; Mismatches 251; Indels 9; Gaps 3;
 QY 4 CCCAATAATCCACACTCCGCTGCTCCGCGCGCGGCATGGCTCCCTCACCTCG 53
 Db 98236 CCCCCCACTCTCTGGCGCTCCCTCTCCGCGGCTCCGCGATGGCTCCCTCGCCTCC 98295
 QY 64 GCGTCACATT---CACTCTCTTCCCGCAGGCTCTCTATCAGGAGCGCATCCGCTCTC 120
 Db 98296 GCGTCCGATACACTGCCTGGTCTTCTCCACCTCTCTCTCCAAAGCCGCGCTCGGCTCC 98355
 QY 121 TCCACCCCCCTGGGCTTCTCCGCGCAGCTCGCGGCTCGGAGCCAGGGGGCGGAGTG 180
 Db 98356 TCCGTCGGATTCTCTCGCGCGCGGTTCCGCGCAGCGCGCGCGGCGGCGCTCCAGG 98415
 QY 181 GCGCGCGGCGCGCGCGGCGCTGCTGGTGTGCGCGCGCGGAGGGGCAAGTTCGAGCGC 240
 Db 98416 GGCACGGGGCGCGCGGGGCTCTGGTGTGCGCGCGGCTAGGGGAAAGTTCGAGCGG 98475
 QY 241 ACCAAACACACGTCACATAGGACCATCGGCCATGTGCACACGGAAGACACACTCTC 300
 Db 98476 ACCAAGCGCAGCTCACATCGGCACCATCGGCCACCTGCAGCACGGAAGACTACGCTG 98535
 QY 301 ACCGCGCGCTCACATGGTGTCTGCGCTCGGTGGCAGCGCGCTTAAGAAGTACGAC 360
 Db 98536 ACGCGGGCTCACCATGGTGTCTGCGCTCGGTGGCGGGAGCGCCCCCAAGAAGTACGAC 98595
 QY 361 GAGATCAGCGCGCCCCCGAGGAGCGCGCGGATATCACCATCAACCGCCACCGTCC 420
 Db 98596 GAGATCAGCGCGCGCGCGAGGAGCGCGCGGCTACCATCAACACCGCCACCGTCC 98655
 QY 421 GAGTACGAGCCGAGACCCGCGCTACGCGACACTGCGACTGCCCGCGCCGCGGACTAT 480
 Db 98656 GAGTACGAGCCGAGACCCGCGCTACGCGCGCTGCGACTGCCCGCGCGCGGACTAT 98715
 QY 481 GTCAAGAATATGATCACCGGCGCTCGCGAGATGACGGTGCATCTCTGCTATCCGGT 540
 Db 98716 GTCAAGAATATGATACCGGCGCGCGAGATGACGGCGCATCTCTGCTGCTCTCGCG 98775


```
QY 1621 TATCTGGTTGCAACTATTGGCTAAGAGTGGCCATCTACTGTTA 1665
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 99850 CATCTGTTTGGCAACTAAATTGGCTATGCAGTGGCCATCTACTGTCA 99894

RESULT 2
AF327413 2050 bp DNA linear PLN 02-DEC-2001
LOCUS AF327413
DEFINITION Oryza sativa translational elongation factor Tu (tufa) gene,
complete cds; nuclear gene for chloroplast product.
ACCESSION AF327413
VERSION AF327413.1 GI:17225493
KEYWORDS Oryza sativa.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
Lee, J.-H. and Kim, J.-K.
1 (bases 1 to 2050)
Chloroplast translational elongation factor Tu gene of rice
Unpublished
2 (bases 1 to 2050)
Lee, J.-H. and Kim, J.-K.
Direct Submission
Submitted (10-DEC-2000) Korea University, 5-1 Anam-Dong, Seong
Buk-Gu, Seoul, Seoul 136-701, South Korea
LOCATION/Qualifiers
1. .2050
/organism="Oryza sativa"
/db_xref="taxon:4530"
<109. >1512
/gene="tufa"
<109. >1512
/gene="tufa"
/product="translational elongation factor Tu"
109. .1512
/gene="tufa"
/codon_start=1
/product="translational elongation factor Tu"
/protein_id="AAL37431.1"
/db_xref="GI:17225494"
/translational="MASLASASATSLVSTSSKPRLGSSVGFSSPARPRRTAAAAA
SRGGRAGLLVRAAKFKPKNVIGTIGHVDHKTLLAALTMVLASVGGSP
KKYDEIDAAPEERARGITINTATVEYETERHYAHVDCPHADYVKNMITGAQMDSA
ILVSGDGPMPQKHEILLAKQVGPRIYVFLNKKDQVDEDELLQLVLELLELJSS
YEDGDEVPVAGSALKALENMANPAIKRGDDEWDGIFSLDSVDNYIPVQRTD
LPFLAVEDVFSITGRGTATRIERTVKGVDIVIGIRTNCTVTVGVEMFQKTM
DDMAGDNVGLLLRGMOKDDIERGMVLAKPASITPHTKFDVAVVVLKDKSGRHSPPF
PGYRPOFVMTDVTGNVTKIMNDKDEAKMCPGDRVKVVELIOPVACEQMRFAI
REGKTVGAGVINTILK"
BASE COUNT 419 a 580 c 550 g 501 t
ORIGIN
Query Match 72.0%; Score 1210.6; DB 8; Length 2050;
Best Local Similarity 84.2%; Pred. No. 2.3e-153;
Matches 1402; Conservative 0; Mismatches 234; Indels 9; Gaps 3;
QY 4 CCCAAATATCCCACTCCCGCTGCTCTCCGCGCCGCCATCGCTCCCTCACCTCG 63
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67 CCCCCCACTCTCTGCGCTCCCTCCCTCCGCGGCTCCGCGCTCCCTCCCTCC 126
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 GGTCCACTT---CACTCTCTCCGCGAGGCTCTCATCCAGGAGCGCATCGTCTC 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 GCTCTCGCATCACCTCCCTGGCTCTTCTCCACCTCTCTCCAAAGCCGCGCTCG 186
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TCCACCCCTCGGCTCTCCGCGAGCTCGCGGCTGCGGAGCCAGGCGGCGCATG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 TCCGTGCGATTCTCTCCGCGCGGGTTCGCGCGACGCGGCGGCGGCGGCTCAGG 246
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GCGCGCGCGCGCGCGGCTGCTGTGTGTGCGCGCGCGGCGGCGGCGGCGGCGG 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 GGCACGGGCGCGCGGCGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 1321 GGTGACCGTATCAAAATGATTTGCTAGCTCATCCAGCCCTGTTGCTGTGTGAGCAGGGTATG 1380
 Db 1387 GGTGACCGTGTCAAGATGTTGTGGAGCTCATCCAGCCCGTGGCTGTGTGAGCAGGAATG 1446
 QY 1381 AGTTTGTCTATCCGTGAGGGTGTAAAGACCGTGTGGTGGCGGTGTCTCATCAACAAATCAT 1440
 Db 1447 AGTTTGTCCATCCGTGAGGGTGTAAAGACCGTGTGGTGGCGGTGTCTCATCAATACGATCTTG 1506
 QY 1441 GAGTAACTGATATACATATACCATCATGAGAAATTTTCTTGTGTACTCAAGCGACAT 1500
 Db 1507 AAGTAACTGATGAGGATATACCATCGTGAGAAATTTTCTTGTGTACTCTTTGGGAAAT 1566
 QY 1501 GCTCCGTAGTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
 Db 1567 GCTCTGTAGTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1626
 QY 1561 ACATTTTTTTTTTGTCAAGTGAATTTTGCATATTTATGATGATGATGATGATGATGATGATG 1620
 Db 1627 GCATTTTTTCTCT--TCAAGTGAATTTTGCATATTTTGTGATATATGATGATGATGATGAT 1680
 QY 1621 TATCTGGTTCACACTCATTTGGCTAAGAGGTGCTCATCTACTGTGA 1665
 Db 1681 CATCTGTTGCACTAATTTGGCTATGCTGTGCACTACTGTCA 1725

AF145053 1678 bp mRNA linear PLN 06-DEC-1999
 Oryza sativa chloroplast translational elongation factor Tu (tufa)
 mRNA, complete cds; nuclear gene for chloroplast product.

AF145053
 AF145053.1 GI:6525064

Oryza sativa (japonica cultivar-group).
 Oryza sativa (japonica cultivar-group).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 1678)
 Lee,J.H., Lee,J.W., Chung,Y.Y., Paek,K.H., Shin,J.S., Yun,C.H. and Kim,J.K.

Cloning and characterization of the chloroplast elongation factor EF-Tu cDNA of Oryza sativa L.
 Mol. Cells 9 (5), 484-490 (1999)

2 (bases 1 to 1678)
 Lee,J.H., Lee,J.W., Paek,K.H., Shin,J.S., Yun,C.H. and Kim,J.K.
 Direct Submission
 Submitted (22-APR-1999) Biology, Korea University, Seong Buk Gu, Anam-Dong, 5-1, Seoul, South Korea

Location/Qualifiers
 1. .1678
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Japonica Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="1"
 1. .1678
 /gene="tufa"
 62. .1465
 /gene="tufa"
 /codon_start=1
 /product="chloroplast translational elongation factor Tu"
 /protein_id="AA15312.1"
 /db_xref="GI:6525065"

translation="MASLASASTSLVSTSSKPRIGSSVGFSSPARPRRTAAAA
 SKGTRRAGLLVMAARKFKERTPHVNIETGRIYHVDPCGHADIVKMMITGAAGSDPA
 KYDEIDAPERRAGITINATVEYETFRIGHVDCPHADIVKMMITGAAGSDPA
 IYVSGADPMPQKEHILLAKQVGPVIVVFNKQDVDEELQLVLEVEVRELLSS
 YLDGDEIVVAGSALKALEMANPAIKRGDDEVDGIFSLIDVDNYIPVPOQRTD
 LPFLIAEDVFSITGRGTAVIERGTGVKGVTDIVIGIRETRNCTVTVGEMFOKTM
 DDAMADNVLLRMOKDDIERGMVLAKPASITPHTKFDAAVYVLKDKSGRHSPPF
 PCYRPOFYRTTDTVGNVVPKIMNDKDEAKMCMFGDRVKVVELLIQVACEQGNRFAL

BASE COUNT 326 a 512 c 485 g 355 t
 ORIGIN
 Query Match 71.4%; Score 1200.6; DB 8; Length 1678;
 Best Local Similarity 83.9%; Pred. No. 5.2e-152;
 Matches 1395; Conservative 0; Mismatches 259; Indels 9; Gaps 3;

QY 4 CCAAATAATCCCACTCCCGCTGCTCCGCGCGCCGCGCTCCCTCACTCCCTCCCTCACTCCG 63
 Db 20 CCCCCCACTCTCTGCGGCTCCCTCTCCGCGGCTCCGCGCTCCCTCCCTCCCTCCCTCCCTCC 79
 QY 64 GGTCCCACTT---CACTCTCTTCCCGCAGCCTCTCTCATCCAGGAGCGCATCCGCTCTC 120
 Db 80 GCCTCCGATCATCCCTCCCTGCTCTCTCCACCTCTCTCTCCAAAGCGCGCTCGGCTCC 139
 QY 121 TCCACCCCTCTGGGCTTCTCCGCGAGCTGCTGCGGCTGCGGAGCCAGCGGCGGAGTGC 180
 Db 140 TCCGTCCGATTTCTCTCGCGCGCGGCTTCCGCGCGCAGCGCGCGGCGGCGGCTCCAAAG 199
 QY 181 GCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCGCGCGCGCGGAGGCGCAAGTTCCGAGCGC 240
 Db 200 GCGAGGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 259
 QY 241 ACCAAACACACGTCAACATAGGCACCATCGGCCATGTCGACACACGAGAAACACACTCTC 300
 Db 260 ACCAAGCGCAGCTCAACATCGGCACCATCGGCCATGTCGACACACGAGAACTACGCTG 319
 QY 301 ACCGCGGCTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 320 ACGCGCGCTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379
 QY 361 GAGTACACGCGCGCGCGCGGAGCGCGCGCTGTCACCACTCAACACCGCCACCGCTC 420
 Db 380 GAGTACGCGCGCGCGCGGAGCGCGCGCTGTCACCACTCAACACCGCCACCGCTC 439
 QY 421 GAGTACGAGACCGAGACCGCGCTACTAGCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 440 GAGTACGAGACCGAGACCGCGCTACTAGCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499
 QY 481 GTCAAGATATGATACCGCGCTGCGCAGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 500 GTCAAGATATGATACCGCGCTGCGCAGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559
 QY 541 GCGGACGCGCGCTGCGCAGACCAAGACACATCTCTCTCGCCAAAGTCTGCTGCTGCTGCTGCT 600
 Db 560 GCGGACGCGCGCTGCGCAGACCAAGACACATCTCTCTCGCCAAAGTCTGCTGCTGCTGCTGCT 619
 QY 601 CCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 620 CCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679
 QY 661 CTCGTGAGCTGCGGCTGCGCAGCTGCTCAGCAACTACGAGTACGAGCGCGGAGGAGTGA 720
 Db 680 CTCGTGAGCTGCGGCTGCGCAGCTGCTCAGCAACTACGAGTACGAGCGCGGAGGAGTGA 739
 QY 721 CCAATGCT 780
 Db 740 CCAATGCT 799
 QY 781 AAGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 800 AAGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 859
 QY 841 TATATTCCAGTCCCGCAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 860 TATATTCCAGTCCCGCAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919
 QY 901 TTTCTCATCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 920 TTTCTCATCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
 QY 961 ATTGGTACACAGTGCATATGCTGCGGAATCGGGACACCCGGAACCTGCACGGTCACTGGT 1020

QY	673	GAGTCCGGAGCTGCTCAGCAACTACGAGTACGACGGCAGACGTAACCAATCGTCGCT	733
Db	1885	GAGTCCGGAGCTTCTTTCCAAGTACGAGTCTCCCGGAGACGATGTTCCGATTATCTCC	1944
QY	733	GGCTCCGCCCTCAAGCGCTCGAGGCTCTCATGGTCAACCCCTGCCCTTGAAGCGGGGAC	792
Db	1945	GGTTCCGCTTGTGCTTTTGAAGCTTTGATGGCCAAACCCGCTCATCAAGCGGGGAA	2004
QY	793	GATGAGTGGCTGACTACATCTTCTCGTTGGTTGATAAAGTGGATTCTATATTCCAGTC	852
Db	2005	AACCAATGGTGTGATAAAATTTACGAGCTCATGGAAGCTGTGGATGACTACATTTCCCATC	2064
QY	853	CCGACAGGAGACTGACCTCCGCTTCTTGCTGCTGTGAAGATGCTTCTTCATCACC	911
Db	2065	CCTCAGCGCAAACTGAACTTCCCTTTTGTCTGGCCATTGAGGATGTTTTCACCATCACC	2124
QY	913	GGTCGTGGTACAGTTGCCACTGGCCGTATAGAGCGTGGCACCGTCAAGATTGGTGACACA	972
Db	2125	GGTCGTGGAAACGTCGCCACCGGAAGGGTTGAGAGAGTACTATTAGGGTTTGGAGAGACT	2184
QY	973	GTCCATATCCTCGGAATCCGGCACACCCCGAACTCCACGGTCACTGGTGTGTGACATGTC	1032
Db	2185	GTGTATATGTTGGTGTGAAGGACACTAGGAACACCACTGTGACTGGGTGGAAATGTTT	2244
QY	1033	CAGAAGCACTGGATGATGCCATGGCCGGGAGACAAATGTTGGGCTGCTGCTCCGTGGTATG	1092
Db	2245	CAGAAGATCTGGATGAGGCTTTGGCTGGGACAAATGTGGCTTGTGCTTAGGGGTATT	2304
QY	1093	CAGAAGATGACATTGAAGAGGCATGTGCTGGCCAAAGCCCTGGCTATTCACACCGCAC	1152
Db	2305	CAGAAACCTGATATTACAGAGGGATGGTTTGGCTAAGCCCTGGAACCACTTACTTCCCTCAC	2364
QY	1153	ACCAAGTTTGAAGGCTGTTGTGTATGCTTTAAGAAGGAAGAGGTGGCCGACACTCACTT	1212
Db	2365	ACCAAGTTTTCGCAATGTGTATGCTGTGTAAGAGGAGAGGTGGAGGCATCTTCCC	2424
QY	1213	TTCTTCCCTGGTTACCGCCACAGTTCTACATCGCGACAACTGATGTGACAGGAGGTG	1272
Db	2425	TTCTTTCGGGTCACGGCCTCAGTTCTACATGAGGACTACCGATGTTACTGGCAAGGTA	2484
QY	1273	ACTACGATTATCAATGACAAAGATGAGAGCGGAGAGATGTGCATGCTTGGTGACCGTATC	1332
Db	2485	ACTGAATTTGAATGACAAAGATGAAGAGTCCAAAGATGGTTATGTCGGCGGACCGTGTT	2544
QY	1333	AAATGATTGTTTCAAGCTCATCCAGCCTTGTGCTGTGAGCAGGATGATGAGTTTGGTATC	1392
Db	2545	AAGCTGGTGTGAATTTATGTGCCTGTGCCTTGTGNACAAAGAAATGAGGTTTGCTATT	2604
QY	1393	CGTGAAGGTGTTAAGACCGTTGGTGGCGGTGCTATCAACAAATCATTTGAGTA	1445
Db	2605	AGAGAAGGAGGCAAACTGTTGGAGCTGGTGTATTCCAATCCATCATTTAGTA	2657

RESULT 5	
GMTUFA	
LOCUS	2313 bp DNA linear PLN 25-JUL-1996
DEFINITION	G.max tufa gene for chloroplast translation elongation factor
	EF-Tu.
ACCESSION	
VERSION	X66062.1 GI:18775
KEYWORDS	translation elongation factor; tufa gene.
SOURCE	soybean.
ORGANISM	Glycine max
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE	1 (bases 1 to 2313)
AUTHORS	Stutz, E.
TITLE	Direct Submission
JOURNAL	Submitted (17-MAY-1992) E. Stutz, Lab. de Biochimie Vegetale, Universite' de Neuchatel, Chantemerle 18, CH-2000 Neuchatel, SWITZERLAND

[illegible]

Db	1116	CCACCGTGGAGTACGAGACCGAGAACCGCCCACTAGCCCACTGCGACTGCCCGCGCCACG	1117
QY	473	CGGACTATGCTCAAGAATATGATCACCGCGCTGCGCAGATGGAGCGGTGCCATCCTCGTGC	532
Db	1176	CTGACTACGTGAANAACATGATCACCGCGCAGCCAGATGGAGCGGCCCATATTATTAGTCG	1235
QY	533	TATCCGGTCCGACGGGCCCATGCCGCAGACAAAGAGCACATCCTCCTCGCCAAACAAG	592
Db	1236	TCTCCGGCGCCGACGGCCCATGCCCAACCAAGAACAACATAATATTAGCGAAACAAG	1295
QY	593	TGGGTGTTCCCAAGATGTTGCTTCCTCAACAAGAGGACATGGTCGACGACGAGGAGC	652
Db	1296	TGGGTGTTCCCAACATGGTTCGTGTTCTTTAAACAAGCAAGCAAAAGTCGAGACGAAGAAG	1355
QY	653	TGCTCGAGTCTGTCAGCTCGAGGTCCGCGAGCTGCTCAGCAACTACGAGTACGACGGCG	712
Db	1356	TTCTCCAACATGTTGAGATAGAGTCCCGACCTTCTGAGCTCCTACGAATTCGCCGGTG	1415
QY	713	ACGAGCTACCAATCTGCGTGGCTCCGCCCTCAAGGGCGCTGAGGCTCTCATGGTCAACC	772
Db	1416	ACGATACCCCATTTGCTCGGGCTCCGCGCTCTTAGCCCTAGAAGCACTCATGGCCCAACC	1475
QY	773	CTGCTTGAAGCGCGGACGATGAGTGGSTCGACTACATCTTCTCGTTGGTTGATAAAG	832
Db	1476	CTGCAATCAACCGCGCGACAACGAGTGGTGCACAAGATTTTCCAACTCATGGACGAGG	1535
QY	833	TGGATTCTTATATCCAGTCCCGCAGAGGCACTGACCTCCCGCTTCTGCTCGCTGTTG	892
Db	1536	TCGACAACCTACATTTCCATCCCCAGCGCCAGACGGACCTCCCTTCTCTCGCGTCG	1595
QY	893	AAGATCTCTTCCATCACCGGTCGTGTACAGTTGGCCACTGGCCGTATAGAGCGTGGCA	952
Db	1596	AAGACGCTTCTCCATCACCGGCGTGGCACCGTGGCCACTGGCCGTAGAGCGTGGCA	1655
QY	953	CCGTCAAGATTGGTGACACAGTGCATATCTGTCGGAATCCGGGACACCCGGAACCTGCACGG	1012
Db	1656	CCATCAAGTAGGGGAAACTGTTGACCTTGTAGGTTTGAGAGAACAAGAACACAACCG	1715
QY	1013	TCACTGGTGTGAGATGTPCCAGAAGACCATGGATGATGCCACTGGCCGGAGACAATGTTG	1072
Db	1716	TCACAGGTGAGAAATGTTCCAGAAGATCTTAGCGAAGCCCTGGCTGGAGACACAGTGG	1775
QY	1073	GGCTGTGCTCCGTGATGTCAGAAGGATGACATTGAAGAGGCAATGGTCTGGCAAGC	1132
Db	1776	GGCTGTGCTTAGAGGGTTCAGAAGACTGACATTCAGAGGGGAATGGTGTGGCTAAAC	1835
QY	1133	CTGGCTCTATCACCGGCACACCAAGTTTGAGGCTGTTGTATGTATGCTTAAGAAAGAAG	1192
Db	1836	CAGGCACGATATAGCGCGACACCAAGTTGCTCAGCGAATTTTATGTTTTGAAGAAGAAG	1895
QY	1193	AGGTTGGCGCACACTCACCTTTCTCCCTGGTTACCGGCCACAGTTCTACATCGCGACAA	1252
Db	1896	AAGTGGTAGGCATTCACCTTTCTTTGCAGGGTATAGCCCTCAGTTTACATGAGGACCA	1955
QY	1253	CTGATGTGACAGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGGAAGATGT	1312
Db	1956	CCGATGTGACTGGGAAGGTTACGTCTATCATGAATGATAAGGATGAGGAGTCCACGATGG	2015
QY	1313	GCATGCCGTGGTACCGTATCAAAATGATTTGCTACGCTCATCCAGCCTGTTGCTTGTGAGC	1372
Db	2016	TGCTGCCCGGTGACCGTGTCAAGATGGTGGTGGAGCTATTGTTCTCTGGCTTGGCAAC	2075
QY	1373	AGGGTATGAGTTTGCTATCCGTGAGGGTGGTAAAGCCGTTGGTGGCGGTGCTCATCAACA	1432
Db	2076	AGGGATGAGTTTGCTATTAGGAAGGTGGGAAGCCGTTGGTGGCTGGTGTATTCCAAAT	2135
QY	1433	AAATCATTTGAGTAAACTGGAATAACAT	1460
Db	2136	CCATTATTGAGTCAAAATCTGTCATT	2163
RESULT	6		
AF234537			

LOCUS	AF234537	1584 bp	mRNA	linear	PLN 15-FEB-2001
DEFINITION	Pelargonium graveolens chloroplast translational elongation factor Tu (tufa) mRNA, complete cds; nuclear gene for chloroplast product.				
ACCESSION	AF234537				
VERSION	AF234537.1	GI:12830554			
KEYWORDS					
SOURCE	Pelargonium graveolens.				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; Geraniales; Geraniaceae; Pelargonium.				
REFERENCE	1 (bases 1 to 1584)				
AUTHORS	Kang,C.J., Lee,M.G., Cho,Y.S., Lee,J.W., Kyung,Y.J., Shin,J.S., Kim,E.S. and Kim,J.K.				
TITLE	Characterization of geranium (Pelargonium graveolens) chloroplast Ep-Tu cDNA				
JOURNAL	Mol. Cells 10 (5), 579-583 (2000)				
MEDLINE	20550852				
PUBMED	11101151				
REFERENCE	2 (bases 1 to 1584)				
AUTHORS	Kang,C.J., Lee,M.G., Cho,Y.S., Lee,J.W., Kyung,Y.J., Shin,J.S., Kim,E.S. and Kim,J.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) Biology, Korea University, Seong Buk Gu, Anam-Dong, S-1, Seoul, Korea				
FEATURES	Location/Qualifiers				
source	1..1584	/organism="Pelargonium graveolens"			
gene	1..1584	/db_xref="taxon:73200"			
CDS	21..1445	/gene="tufa"			
		/codon_start=1			
		/product="chloroplast translational elongation factor Tu"			
		/protein_id="AAK08141.1"			
		/db_xref="GI:12830555"			
		/translation="MALPVAISAKLLCPHASPSPSLTSLSTSEFSPKSIPTLNLSS TSLMSTAALSTRRRRTVRAARKFERKKPHVNIT IGVDHGKTTLTAALTMAIA SLGNSAPKYDEIDAAPERARGITINTATVEYETENRHYAVDPCGHADVKNMTIG AAMDGAIIYVGSGADGPMPQTKHEILLAKQGVPMNVFLNKQDVODELELLEVELE VRELSAYEPFGDPVPPIISSGALLALEALMAPAIKRGENOVFLNRDTKRSVTIVG LPQKDLPFLAAGDNVGLLRGVQKEDATRGVLAKPGTIPHTKFSAIYVYLKKEGG EMFOKITDELADGNVGLLRGVQKEDATRGVLAKPGTIPHTKFSAIYVYLKKEGG GRHSFPFAYRPQFVRFTDVTGKVATIMNDKEESKMVMPEGDRVKMVVELILPACGE QGMFAIRREGSKTVGAVISSIIE"			
BASE COUNT	357 a 438 c 396 g 393 t				
ORIGIN					
Query Match	46.0%;	Score 774;	DB 8;	Length 1584;	
Best Local Similarity	71.9%;	Pred. No. 9.3e-95;			
Matches 1011;	Conservative 0;	Mismatches 395;	Indels 0;	Gaps 0;	
Qy	41 CGGCATCGCCTCCCTCACTCGGCGTCACATTCACTCTCTCTCCGCGAGGCCTCTCAT 100				
Dd					
Db	40 CAGCGCAAGCTCTTATGCCACACAGCGCTCTCCTCTCCTCTCCACTCTCTCACTCCCCT 99				
Qy	101 CCAGGACCGGATCGTGCTCTCCACCCCCCTGGGCTTCTCCGCGAGCCTCGCGCGCTGC 160				
Dd					
Db	100 CAACCTCTCTTCCACTAACCCCTCAAATCTCACCCCAACCAATCTCTCTCCACTCC 159				
Qy	161 GGACGACGGGGCGGCGAGTTGGGCGCGCGCGCGCGCGCGCTGTGTGGTGC GCGGG 220				
Dd					
Db	160 TCACCCCTCCACCGCGCCATTTCACGCGCGCGCGCGCGCTTACCCTCCGCGCG 219				
Qy	221 CGAGGGCAAGTTCGAGCGCACCAACAACACAGCTCAACATAGGCACCATCGGCCATGCG 280				
Dd					
Db	220 CCCGCGCAAGTTCGAGCGCAAGAAGCCCCATGTCAACATCGGCACCATCGGCCACGCTG 279				
Qy	281 ACCACGGAAGACCACTCTCACCGCGCGCTCACCATGGTGTCCGCTCCGCTGCGTGCCA 340				
Dd					
Db	280 ACCACGGAAGACCACTCTCACCGCGCGCTCACCATGGTGTCCGCTCCGCTGCGTGCCA 339				

200

[illegible]

173

Db 436 CTCGGTAACAGCGCCCTAAAAAGTACGAGAAATGACGCTGCTCCGAGAGCGTGCA 495
 QY 391 CGCGGTATCACCATCAACACCGCCGCTGAGTACGAGACGAGACCGCCACTACCA 450
 Db 496 CGTGGATTAACAATTAACACTGCAACCGTTGAGTACGAGACTGAACACTGTCACATCGCT 555
 QY 451 CAGTCGACTGCCCCCGCCACCGCTACTGTCAAGAATATGATCACCAGCGCTCGCGAG 510
 Db 556 CATGTGATGTTCCCGGTACGCTGATTAAGTCAAGAATATGATCATTGTCGCGCTCA 615
 QY 511 ATGAGCGGTGCTATCTCTGCTGATTCGCTGCGGAGCGGCCATGCGGAGACCAAGAG 570
 Db 616 ATGAGCGGCTATCTCTGCTGCTGCGGCGGATGTTGCCATGCTCAAACTAAAGAA 675
 QY 571 CACATCTCTCTGCCAAGCAAGTCTGCTTCCCAAGATGCTTCTCTCAACAGAG 630
 Db 676 CACATCTCTCTGCCAAGCAAGTCTGCTTCCCAAGATGCTTCTCTCAACAGAG 735
 QY 631 GACATGCTGAGACGAGGAGTCTGCTGAGCTCTGCTGAGCTCGAGCTCGCGAGCTGCTC 690
 Db 736 GACCAAGTGGATGATGAGGAGCTCTGAGCTTGTGAGCTTGAAGTTCGGGAGCTTCTC 795
 QY 691 AGCAATACGATACAGCGGCGAGCAGCTACCAATGCTGCTGCTGCTCCGCTCAAGGG 750
 Db 796 TCATCTTATGAGTTTCTCTGCGATGACATTCGATTTCTTCTGCTGCGGCTCTTGGCC 855
 QY 751 CTGAGGCTCTCATGCTCAACCTGCTTGAAGCGCGGAGCAGTACGAGTGGTCTGACTAC 810
 Db 856 TTGGAAGCATTAATGCAACCCCTACTCTTAACGTTGGAATTAACGAGTGGGTGATAG 915
 QY 811 ATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
 Db 916 ATTATCAGCTTATGATGAAGTGTGATAATATATCCCAATCCCTCAACGTCAGACGAA 975
 QY 871 CTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
 Db 976 CTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035
 QY 931 ACTGGCGGTATAGAGCGTGGCAGCGTCAAGTGTGTGACACAGTTCGATCTCGGAAATC 990
 Db 1036 ACTGGGAGATTTGAAGAGCGGTGTTAAAGTTGGTGTGATGCTGCTGCTGCTGCTGCTG 1095
 QY 991 CGGACACCCCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
 Db 1096 AGGGAACAAGGAACACTACCTGCTGAGCGGTGTAGAAATGTTCCAAAGATTTTGGATGAT 1155
 QY 1051 GCCATGCGCGGAGACATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
 Db 1156 GCTATGCGCGGTGACAAATGCGGTTGTTGTTGAGAGGTATTCAGAAGATTTGACATTCAA 1215
 QY 1111 AGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1170
 Db 1216 AGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1275
 QY 1171 GTGATGCT 1230
 Db 1276 GTATGCT 1335
 QY 1231 CCACGCTTCTACATCGCGGACAACTGATGTCAGAGGAGTGTGACTACGATTAATGAATGAC 1290
 Db 1336 CCTCAATTTTACATGAGGAGCAGAGATGTAACCTGGAAGATTTACTTCTATTATGAATGAT 1395
 QY 1291 AAGGATGAGGAGGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350
 Db 1396 AAGGATGAGGAGGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1455
 QY 1351 ATCCAGCT 1410
 Db 1456 ATTTGTTCCCGTGTATCAACAGGAATGAGTTCGCTATTAGAGAGGAGGAGGAGAC 1515
 QY 1411 GTTGGCGCGGTGTATCAACAAATCAATTTAGTAAA 1447
 Db 1516 GTTGGAGCT 1552

RESULT 8

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU

LOCUS

DEFINITION

TOBTEFTU


```
Db 827 ATTACCGTGTGCTGCCAAATGAGCGGGCTATTCTTCTGTTTTCAGGTGCTGATGGTCCC 886
Qy 553 ATGCGCAGACAAAGAGCAGCATCTCTCGCCAAAGCAAGTCGGTGTCCCAAGATCGTT 612
Db 887 ATGCCACAGCTAGGAACATTTTCTGCTAAGCAAGTAGTGTCCCAACATCGTT 946
Qy 613 GTCTTCTTCAAAAGAGCAGCATCTCGAGCAGCAGGAGCTGTCTCGAGCTCGTCGCTC 672
Db 947 GTGTTCTTGAACAAACAGATCAGTGTGATGATGAGGAGCTGCTCAGCTTGTGAGTTG 1006
Qy 673 GAGTCCCGAGCTGCTCAGCACTAGAGTACAGCAGCGGACGAGCTACCAATCGTGGCT 732
Db 1007 GAGGTAAGAGAGTATTGTCAAGTTATGAGTTCTCGTGTGATGATATCTTATTTCT 1066
Qy 733 GGCTCCGCTTCAAGCGCTCGAGGCTCTCATGTGCTCAACCTGCTTGAAGCGCGGAC 792
Db 1067 GGCTCTGCTTTGGCTTTAGAGGCTTTGATGGCTAATCTAGTATTAAGAGAGTGAA 1126
Qy 793 GATGAGTGGTGCAGTACATCTTCTGCTGTTGATGAAGTGGATTCCTATATTCAGTC 852
Db 1127 AATCAATGGTGTGAAGATATATGATGTGATGAGCGCTGTGATAGTTATATTCCTATT 1186
Qy 853 CCGCAGAGGACAGTACCTCCCTCTTCTGCTGCTGTTGAAGATGCTTCTCCATCAC 912
Db 1187 CCAGTTAGGCAACTGAATTCCTCTTCTGATGGCTATGAAGATGTTTTCATTA 1246
Qy 913 GGTGCTGTACAGTGGCAGTATAGAGCGTGGCAGCGTCAAGATTTGGTGACACA 972
Db 1247 GGTAGAGTACTGTGGCAGCGGAGGTAGAGAGGAGTGTAGGATGGAGACACT 1306
Qy 973 GTCGATATCGGAATCCGGGACACCCGGAACCTGCAAGTGCAGTGTGTGAGATGTT 1032
Db 1307 GTTGATATTTAGTGTGTTGAAGGACACTAGGAGTACTACCTGAGCGGTGTGAGATGTT 1366
Qy 1033 CAGAAGACCATGATGATGCTGCGCGGAGACAATGCTGCGTGTGCTGCTGCTGATG 1092
Db 1367 CAAAGATTTTGGATGAAGCAATGGCTGGAGACAATGCGGTTGTTGTTGAGAGGTATT 1426
Qy 1093 CAGAAGATGATGATGAAGAGGATGTTGCTGGCAAGGCTGCTCTATCACACCGCAC 1152
Db 1427 CAAAGATTTGATATTCAGAGAGGAATGTTGTTGCAAAACCCGGAACAATACCCCTCAC 1486
Qy 1153 ACCAAGTTGAGCTGTTGTTATGCTTTAAGNAGGAAGAGGTCGCCACACTCACCT 1212
Db 1487 ACCAAGTTTGAAGTATTTGTTATGTTTGAAGAGGAGGAGGTTGAGTGCATTCCTCC 1546
Qy 1213 TTCTTCTCTGTTACCGCCACAGTCTTACATGCGGCAACTGATGTCAGGAGGAGTGTG 1272
Db 1547 TTCCTTTCAGGTCACAGGCTCAGTTTACATGAGGACTACTGATGTCGCGGAAGGTT 1606
Qy 1273 ACTAGGATTAATGATCAAGGATGAGGAGGCGGAAGATGTCATGCTGTGTCACCGTATC 1332
Db 1607 ACTTCCATTAGCACTAAGAGGAGAGGAATCAAGATGCTATGCTGCTGCTGATCGTGTG 1666
Qy 1333 AAAATGATGTTACAGCTCATCCAGCTTGTGCTGTCAGCAGGATGAGGTTGCTATC 1392
Db 1667 AACTTGGTGTGAGCTCATTTATCCGGTGGCTTGTGAGCAAGGATGAGATTTGCCATC 1726
Qy 1393 CGTGAGGTTGTAAGACCGTTGTTGCGCGTGTGTCATCAACAAGATCATTTAGTAACTGGA 1452
Db 1727 AGGAGAGGAGGAAGACTGTTGGAGCTGTTGTCATTCAGAAATATTCAGAGTATCAACC 1786
Qy 1453 TATAACATATCCA 1465
Db 1787 AATAGCGGAGCAA 1799

RESULT 9
TOBTUFA2
LOCUS
DEFINITION Nicotiana sylvestris tufa gene for chloroplast elongation factor
Tua, complete cds.
ACCESSION D11469
```

D11469.1 GI:459238
g-binding protein; elongation factor; nuclear-encoded chloroplast
elongation factor TuA(EF-TuA).
Nicotiana sylvestris DNA, clone_lib:lambda Dash.
Nicotiana sylvestris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 2667)
Sugita,M., Murayama,Y. and Sugiyama,M.
Structure and differential expression of two distinct genes
encoding the chloroplast elongation factor Tu in tobacco
Curr. Genet. 25 (2), 164-168 (1994)
94373864
2 (bases 1 to 2667)
Sugiyama,M.
Direct Submission
Submitted (19-JUN-1992) Masahiro Sugiyama, Nagoya University, Center
for Gene Research (Tel:052-789-3081, Fax:052-789-3081)
Location/Qualifiers
1..2667
/organism="Nicotiana sylvestris"
/db_xref="taxon:4096"
/clone_lib="lambda Dash"
257..265
CAAT_signal
288..294
TATA_signal
331..337
TATA_signal
352..357
misc_feature
388
/note="transcription start site"
551
/note="transcription start site"
668..2104
/gene="tufa"
668..2104
/gene="tufa"
/codon_start=1
/product="chloroplast elongation factor TuA(EF-TuA)"
/protein_id="BAA0207.1"
/db_xref="GI:459239"
/translation="MAISAAATSTKLYSSNTNPLPSSTKPSKLILSSSFTPNP
STLHSPATSSATTHRRRTVRAARGKPKPHVNIQTIGHVHGKTLTAALT
MALASGNAPKYEIDDAPEERARGITINTATVEYETENRHHVHDCPGHADYVN
MITGAAMDGAIVCSGADGMPQTKHLLAKOVGNVNVFLNKQDQVDELLQL
VELEVRELLSYEPFGDDIPTISGALLAALMANPSIKRGENQWDKIYELMDVD
SYIPVPTLPEFLMAIEDISITGRTATGRTATGRTATGRTATGRTATGRTATG
VTGEMFOKILDEAMAGNVGLLRGIQIDIORGMVLAKPGTITPTKFAIVVILK
KEGRHSFESGYRPOFYMTTDTGKVTISITTDKGEESKMWMPGDRVNLVVELIMP
VACGGHRAFAIRBGGKTVGAGVIOKIE"
transit_peptide 668..877
/gene="tufa"
878..2101
/gene="tufa"
/product="chloroplast elongation factor TuA(EF-TuA)"
BASE COUNT 758 a 528 c 533 g 848 t
ORIGIN
Query Match 41.7%; Score 700.2; DB 8; Length 2667;
Best Local Similarity 71.9%; Pred. No. 7e-85;
Matches 915; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 193 CGCGGGCGCTGCTGGTGTGCGCGGAGGGGCAAGATTGCGAGCGCAACACAC 252
Db 851 CGCCACCGCGCTTTACTGTCGCGCTGCCCTGGCAATTCGAGCGGAAAAACCCCAT 910
Qy 253 GTCAACATAGCACCATCGGCATGTCGACACGAAAGACCACTCTCACCGCGGCTC 312
Db 911 GTCAACATCGGTACTATTGGCCATGTTGACACGGAAGACTACTCTCACCGCTCTTA 970
Qy 313 ACCATGGTCTCGCTCGCTCGCGCGGCGGCTTAAGAGTACGAGATCGACGCC 372
Db 971 ACCATGGTCTTCTCTATGGAACATCCGCCCTTAAGAAATACGACGAATTCATGCC 1030

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CAAT_signal
CAAT_signal
TATA_signal
TATA_signal
misc_feature
misc_feature
gene
CDS
transit_peptide 668..877
mat_peptide
BASE COUNT 758 a 528 c 533 g 848 t
ORIGIN
Query Match 41.7%; Score 700.2; DB 8; Length 2667;
Best Local Similarity 71.9%; Pred. No. 7e-85;
Matches 915; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 193 CGCGGGCGCTGCTGGTGTGCGCGGAGGGGCAAGATTGCGAGCGCAACACAC 252
Db 851 CGCCACCGCGCTTTACTGTCGCGCTGCCCTGGCAATTCGAGCGGAAAAACCCCAT 910
Qy 253 GTCAACATAGCACCATCGGCATGTCGACACGAAAGACCACTCTCACCGCGGCTC 312
Db 911 GTCAACATCGGTACTATTGGCCATGTTGACACGGAAGACTACTCTCACCGCTCTTA 970
Qy 313 ACCATGGTCTCGCTCGCTCGCGCGGCGGCTTAAGAGTACGAGATCGACGCC 372
Db 971 ACCATGGTCTTCTCTATGGAACATCCGCCCTTAAGAAATACGACGAATTCATGCC 1030

3'UTR
BASE COUNT 398 a 353 c 393 g 519 t
ORIGIN

Query Match 40.8%; Score 685.6; DB 8; Length 1663;
Best Local Similarity 72.2%; Pred. No. 6.8e-83;
Matches 892; Conservative 0; Mismatches 344; Indels 0; Gaps 0;

TGAOMDGAIIIVSGADGPMPTKEHILLAKOVGPDVFLNKEDVDVAELLEIVE
LEVRELLSSVEFNGDDIPISGSALLAVETLLENPKVKRGDNKWDKIYELMDVDDY
IPPOKTELPFLAVEDVFSITGRGTVAIGRVERGTGVKGVSETDLVLGURETSYTVT
GVEMFKIILDEALAGNVLGIIQKADIORQWVLAKPGSITPHTKFBAIYVLKKE
EGGRHSFPFAGYRQPYMRITDVTGKTKIMNDKDESKWMPGDRVKIIVWELLVPVA
CEGMRFAIREGSKTVGAGVIGFILE"
1498..1663

Qy 211 GTGGCGGGGAGGGGAGATTGCGAGCGCACCAAAACACACGCTCAACATAGGACCATC 270
Dy 262 GTGGCGGGCGCTGCTGGAAGATTGAGAGGAAGAAGCTCATGTCAACATCGGAACATC 321
Qy 271 GGCGATGTCGACCGAGAAAGACCACTCTCACCGCGCGCTCACCATGGTGGCTCGCTCC 330
Dy 322 GGTCAATGTTGACATGGGAACACTCTTAACCGCACCTCTAACCATGGCTCTCGCTCC 381
Qy 331 GTGGTGGCGGCGCTTAAGAAGTACGACGAGATCGACCGCGCGCTCACCATGGTGGCTCGCTCC 390
Dy 382 ATTGGTCCAGCGCTCAAAAGTACGAGAGATTGACGCTGCGCGGAGGAGAGCT 441
Qy 391 CGCGGTATCACCATCAACACCGCCACCGCTCGAGTACGAGCGGAGACCGCCGCTAGGCA 450
Dy 442 CGTGTATCAAAATCAACACTGCTACTGTTGAGTACGAGAGTGAAGATCGCTACCTAGCT 501
Qy 451 CACGTGCACTCCCGCGCGCGCTGCTGCTCAAGATATGATCAGCGCGCTGCGCGAG 510
Dy 502 CACGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
Qy 511 ATGGAGCGTGCATCTCTGCTGCTATCGCGTCCGAGCGGCGGCGGCGGCGGCGGCGGCG 570
Dy 562 ATGGAGCGGAGTATCTCTGTTTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 621
Qy 571 CACATCTCTCTGCGCGAGCGAGTGGTGTCCCAAGATCGTGTCTCTCTCAACAAGAG 630
Dy 622 CATATCTCTTGGCTTAAGCAGGTTGGTGTCTCTGATATGTTGTTCTCTTAACAAGAG 681
Qy 631 GACATGCTGCGAGCGAGGAGTCTCTGAGTCTGCTGAGTCTGAGTCTGCGGCGGCGGCTCTC 690
Dy 682 GATCAAGTAGATGATGACGAGTGTCTGAGCTGCTGAGCTGAGTCTGAGTCTGAGCTCTC 741
Qy 691 AGCAACTAGAGTACGAGCGGCGAGCTACCAATCGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
Dy 742 TCGTCTTATGAATTAACGGTGAATATCCGATATCTCTGCTGCTGCTGCTGCTGCTGCT 801
Qy 751 CTCGAGGCTCTCATGCTCAACCCCTGCTGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCTAC 810
Dy 802 GTTGAGACTCTTACTGAGAACTCTTAAGTTAAGAGAGGTGATACAAATGGGTAGATAAG 861
Qy 811 ATCTTCTCTGTTGTTGATAAGTGGATTCCTATATTCAGTCCCGCGAGCGAGCAGTAC 870
Dy 862 ATTTATGAACATATGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 921
Qy 871 CTCGCGTCTGCTGCTGTTGAGATGCTCTCTCCATCCGCTGCTGCTGCTGCTGCTGCTGCT 930
Dy 922 TTGCGCATCTCTGCTGCTGTTGAGAGTGTCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 981
Qy 931 ACTGGCGCTATAGAGCGTGGCAGCGCTCAAGATTGCTGACAGTCTGATGCTGCTGCTGCTGCT 990
Dy 982 ACAGGCGCTGCTGAGAGAGGTACGTTAAGTTAGGAGAGAGCTGTAGATTAGTGGTTG 1041
Qy 991 CGGAGACCGCGAAGTCCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
Dy 1042 AGGAGACTAGGAGTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
Qy 1051 GCCATGCGCGGAGCAATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
Dy 1102 GCTTTAGCTGGTGACAATGATAGGTTGTTGCTTAGGGGTATTTCAAAAGGCTGATATCGAG 1161

Qy 1111 AGAGGATGCTGCTGCAAGGCTGCTCTATCACACCGCACCAAGTTTTCAGGCTGTT 1170
Dy 1162 AGAGGATGCTGTTTATTAAGCGGGATCGATTACTCCACATACCAAGTTTTCAGCAAT 1221
Qy 1171 GTGATGCTGCTTAAAGAAAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1230
Dy 1222 ATCTATGCTTGAAGAAAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1281
Qy 1231 CCACAGTTTACATCGGCACTGATGACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290
Dy 1282 CCTCAGTTTACATGAGGAGGCTGATGTTAGGGTAAAGTGAAGAGTATCATGACGAC 1341
Qy 1291 AAGGATGAGGAGGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350
Dy 1342 AAGATGAGAGTCGAGGATGTTATGCGCGGTCGAGTGAAGATGTTGTTGAGCTT 1401
Qy 1351 ATCCAGCTGTTGCTGTTGAGGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1410
Dy 1402 ATTTGCGGCTGCTGTTGTAACAAGGATGAGTTTGTCTATCAGAGAAGGAGGAAGACT 1461
Qy 1411 GTTGGTGGCGGCTGCTATCAACAAATCATGAGTAA 1446
Dy 1462 GTTGGTGGCGGATTTATGGGAGGATCCTCGAATGA 1497

RESULT 12
AF410329
LOCUS Arabidopsis thaliana AT4g20360/F9f13_10 mRNA, complete cds.
DEFINITION AF410329
ACCESSION AF410329.1 GI:15294275
VERSION
KEYWORDS FLI-CDNA
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1669)
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1669)
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (13-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
COMMENT
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,

Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologos,A., and Ecker,J.R.

Chen,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

1..1669

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="4"

/clone="RAF106-16-G16(R13435)"

/note="ecotype: Columbia"

1..64

5'UTR

CDS

65..1495

/note="translation elongation factor EF-Tu precursor,
chloroplast"

/codon_start=1

/product="At4g20360/F9F13_10"

/protein_id="RAK95315.1"

/db_xref="GI:15294276"

/translation="MAISPAACSSSRILCSYSSPSPSLCPAISTSGKLTLLSS
FLPSYLLTTSASQSRSTFRAARGKFERKPHVNIQITIGHVDHGKTLTAALPMA
LASISGVAKYDEIDAAPEERAGITINTAVEYETENRHYAHVPCDGHADYVKMI
TGAQMDGAILVVGADGPMPOKHEHLLAKOVGPDVDFVFNKEDQDDAELELVE
LEVRLLSSFEFGDDIPILSGSALLAVETLENPKVKGDKNDKIELMDAVDDY
IPIPORTELPIFLAVEDVFSITGRGTAVRVERGTVKVGETVDLVGLRFSRYVT
GVEMFKILDEALAGNVGLLRIQKADIQRMVLAKPGSITPHKFAIIVLKE
EGHRSFFAGYPOFYMRITDVTGKVTIMNDKDESKWMPGDRVKIVVELIPVA
CEQMRFARIEGGKTVGAVIGTILE"

1496..1669

3'UTR

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 40.8%; Score 685.6; DB 8; Length 1669;

Matches 892; Conservative 0; Mismatches 344; Indels 0; Gaps 0;

QY	211	GTGCGCGCGGCGGCAAGTTCGACGCGCACCAACACACGTCACATAGGACCATC	270
DB	260	GTCCGCGCGGCTGCGAAAGTTCGAGAGGAAGGCTCATGCAACATCGGAACATC	319
QY	271	GGCATGTCGACACGCGGCAAGTTCGACGCGCACCAACACGTCACATAGGACCATC	330
DB	320	GGTCACTTGACCATGCGGAAACTACTTTAACCGGAGCTTAACCATGCTCGGCTTC	379
QY	331	GTGCGTGGCGGCGGCTAAGAGTACGAGGATCGACGCGCGGCGGCGGAGGAGCT	390
DB	380	ATTGGTTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	439
QY	391	CGCGGTATACCATCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT	450
DB	440	CGTGGTATCAATCAACACTGCTACTGTGTAGTACGAGACTGAGAATCGTCACTACGCT	499
QY	451	CAGTCGACTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	510
DB	500	CAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	559
QY	511	ATGAGCGGTGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	570
DB	560	ATGAGCGGAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	619
QY	571	CACATCTCTCGGCAAGCAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	630
DB	620	CATATCTCTTGGCTAAGCAGGTTGGTGTCTGATGTTGGTGTCTTAAACAAAGAG	679
QY	631	GACATGCTGACGACGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	690
DB	680	GATCAAGTAGATGATGACAGAGTTGCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT	739
QY	691	AGCAACTAGGATGACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	750

Db	740	TCGCTTATGAATTAACGGTGATGATATCCGATTATCTGCTGCTCTTTAGCC	799
QY	751	CTCAGGCTCTCATGTCACCCCTTGAAGCGCGGACGATGAGTGGTGCAGTAC	810
Db	800	GTGAGACTCTTACTAGAAATCCTAAGGTTAAGAGAGGTGATAACAAATGGTAGATAAG	859
QY	811	ATCTTCTCGTGGTGTATGAAGTTCCTATATTCCTATATTCCTGCGGACGAGAGTGC	870
Db	860	ATTTATGAACCTTATGATGCTGTGTGATGATATCATCCCTATCCCTCAGAGACAACTGAA	919
QY	871	CTCCCGTCTTGTCTGCTGTTGAAGATGCTTCTTCCATCACCGGTCCTGCTGATGCTG	930
Db	920	TTGCCATCTTGTGTAGCTGTTGAGGATGCTTCTCTATCACTGAGCTGCTGCTGCTGCT	979
QY	931	ACTGGCGGTATAGAGCTGGCACCCTGCAAGATTGGTGACACAGTTCGTCGGAATC	990
Db	980	ACAGGCGGTGTCGAGAGAGGTACGGTTAGGTAGGAGAGCTGTAGATTAGTGGGTTG	1039
QY	991	CGGACACCCCGAACTGCACGGTCACTGGTGTGAGATGTTCCAGAGACCATGGATGAT	1050
Db	1040	AGGAGACTAGGATTTACACTGTCACTGGGTTGAAATGTTTCAGAAAGATTCTTGATGAG	1099
QY	1051	GCATGCGCGGAGACAATGTTGGCTGCTGCTCGGTGATGCGAGAGGATGACATTGAA	1110
Db	1100	GCTTTAGCTGGTGACAAATGAGGGTGTGCTTAGGGGTATTCAAAGGCTGATTTAG	1159
QY	1111	AGAGGATGCTGCTGCAAGCTGCTCTATCACACCCACACCAAGTTTGGGCTGT	1170
Db	1160	AGAGGATGCTGCTGCAAGCTGCTCTATCACACCCACACCAAGTTTGGGCTGT	1219
QY	1171	GTGATGCTTAAAGGAAGAGGGTGGCGACACTCACTTCTTCCCTGGTTACCGC	1230
Db	1220	ATCTATGCTTGAAGAGAGGAGGTTGAAGGCTTCTTCCCTGGTTACCGC	1279
QY	1231	CCACAGTCTTACATGCGGACAATGATGTGACAGGAGTGTGACTACGATTATGATGAC	1290
Db	1280	CCTCAGTCTTACATGAGGAGGAGTGTGATGAGGAGTGTGACTACGATTATGATGAC	1339
QY	1291	AAGGATGAGGAGGAGGAGTGTGATGAGGAGTGTGACTACGATTATGATGAC	1350
Db	1340	AAAGATGAGGAGGAGGAGTGTGATGAGGAGTGTGACTACGATTATGATGAC	1399
QY	1351	ATCCAGCTGTTGCTGTGACAGGATGATGAGGTTGCTATCCGTTGAGGTTGTTAGGAC	1410
Db	1400	ATTGTCGGTGGCTGTGTAACAGGAGTGTGCTATCAGAGAGGAGGAGGAACT	1459
QY	1411	GTGTCGGTGTGCTATCAACAAATCATTTAGTAA	1466
Db	1460	GTGTCGGTGTGCTATTTGGGAGGAGTCTCGAATGA	1495

RESULT 13

AV074355

LOCUS

DEFINITION

precursor, chloroplast (At4g20360) mRNA, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana.

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1688)

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,

Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,

Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,

Satou,M., Seki,M., Shinozaki,K., Southwick,A., Shinozaki,K.,

Davis,R.W., Ecker,J.R. and Theologis,A.

AY074355 1688 bp mRNA linear PLN 26-APR-2002
Arabidopsis thaliana putative translation elongation factor EF-Tu
precursor, chloroplast (At4g20360) mRNA, complete cds.

AY074355.1 GI:18377802

FLI.CDNA

Arabidopsis thaliana.

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1688)

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,

Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,

Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,

Satou,M., Seki,M., Shinozaki,K., Southwick,A., Shinozaki,K.,

Davis,R.W., Ecker,J.R. and Theologis,A.

	DVEKLLIKIGLKIPATNTAETRAAAQTHGISNPKDKKKKKQEISKRTKLTNAHLPELLE
exon	ONLNGSSSRN" complement(1281. .1631) /gene="F9f13.2" /number=1 complement(1632. .1958) /gene="F9f13.2" /number=1 complement(1959. .2075) /gene="F9f13.2" /number=2 complement(2076. .2167) /gene="F9f13.2" /number=2 complement(2168. .2299) /gene="F9f13.2" /number=3 complement(2300. .2376) /gene="F9f13.2" /number=3 complement(2377. .2637) /gene="F9f13.2" /number=4 3609. .6346 /gene="F9f13.4" join(3609. .3664, 3783. .3902, 4023. .4143, 4323. .4436, 4561. .4670, 4763. .5173, 5213. .5372, 5484. .5669, 5813. .6017, 6246. .6346) /gene="F9f13.4"
intron	/note="weak similarity to transcription factor IIE alpha-xenopus laevis, PIR2:S26646" /codon_start=-1 /product="putative protein" /protein_id="CAB52822.1" /db_xref="GI:5738379"
exon	/translation="MDKSITVVRKVTVLPFPVKLVRLIVRFDNYTSPESNQOKSVK NVGSAVVLDALTRRWVREEDLAKEVRKAELKLRIHFEOKFVMRYHRKETAK RAKMYSVAVGTTDGRAEDNVAFPTHSYCYCLDYAOIYDIVRVKLRHKKFKDELEDNR NTVOEGCPKNCKBYNWDALRLISMEDDSFHCCNGELYMECNKLI SEEVVDGRGN ARROREGKPVKWLODELVCYCFCGVNESMKETOIDLPSLKLVHRSKFMKMSRYMI KFRPRILKYPMLEICKTKQDCGCSCVCDTI GNKIYKDLPFPFEPFAWEARAKAAA RENGDNPDDPSNLGGYGSTPMPFIGETFKVNINEGNEDEDVTSGDGSDSLSUKMLPPNM IKQGKMLETQRGMOEANVDGEAAKLSDDKSDKEYIAKAYAAIMEQQKLAALKINEQ ESAGESTTTDI EASATTYSDRQVMKSKREEEDEVEEGASVAANGNKYKVDLNAV EEAEKKGDDGEDDDIWEEG"
exon	3609. .3664 /gene="F9f13.4"
intron	/number=1 3665. .3782 /gene="F9f13.4"
exon	/number=1 3783. .3902 /gene="F9f13.4"
intron	/number=2 3903. .4022 /gene="F9f13.4"
exon	/number=2 4023. .4143 /gene="F9f13.4"
intron	/number=3 4144. .4322 /gene="F9f13.4"
exon	/number=3 4323. .4436 /gene="F9f13.4"
intron	/number=4 4437. .4560 /gene="F9f13.4"
exon	/number=4 4561. .4670 /gene="F9f13.4"
intron	/number=5 4671. .4762

/gene="F9F13.4"
/number=5
4763. .5173
/gene="F9F13.4"
/number=6
5174. .5212
/gene="F9F13.4"
/number=6
5213. .5372
/gene="F9F13.4"
/number=7
5373. .5483
/gene="F9F13.4"
/number=7
5484. .5669
/gene="F9F13.4"
/number=8
5670. .5812
/gene="F9F13.4"
/number=8
5813. .6017
/gene="F9F13.4"
/number=9
6018. .6245
/gene="F9F13.4"
/number=9
6246. .6346
/gene="F9F13.4"
/number=10
6989. .8331
/gene="F9F13.6"
join(6989. .7084, 7162. .7222, 7316. .7354, 7454. .7566,
7703. .7925, 8033. .8079, 8185. .8331)
/gene="F9F13.6"
/note="similarity to hypothetical protein - Caenorhabditis
elegans, SPTREMBL:Q17527"
/codon_start=1
/product="putative protein"
/protein_id="CAB52823.1"
/db_xref="GI:5738380"
/translation="MELEFRVGLTPTVYIPGFTIDERTQTLNHHYGAASKWKL
KNRLQNGMVEHKLVPQELPWLTKITAEIHSSGLFPSSAIHNLHNEYHPDGI
MPHQDPAYPPVVAIISLGSPPVYMDFTPHLRISGDGYISKQSPCAESCAPDSFS
VLLMQSLIFIKDDAYSDFLHGISDPTQCYNOVNEALAYSNEEDSRKDGDKIFH
RDOTRYSLTCRLVPKVRKNLFR"
6989. .7084
/gene="F9F13.6"
/number=1
7085. .7161
/gene="F9F13.6"
/number=1
7162. .7222
/gene="F9F13.6"
/number=2
7223. .7315
/gene="F9F13.6"
/number=2
7316. .7354
/gene="F9F13.6"
/number=3
7355. .7453
/gene="F9F13.6"
/number=3
7454. .7566
/gene="F9F13.6"
/number=4
7567. .7702
/gene="F9F13.6"
/number=4
7703. .7925

Query Match 40.8%; Score 685.6; DB 8; Length 109936;
Best Local Similarity 72.2%; Pred. No. 4e-83;

	Matches	892;	Conservative	0;	Mismatches	344;	Indels	0;	Gaps	0;
QY	211	GTCCGGCGGGGAGGGGCGAAGTTTCGAGCGGACCAACACACAGCTCAACATAGGCACCAATC	270							
DB	8829	GTCCGGCGGGGAGGGGCGAAGTTTCGAGAGGAAGAGCCCTCATGTCAACATCGGAACCAATC	8888							
QY	271	GGCCATGTGACACACGGAAGACCACTCTCACCGCGCGCTACCATGTGTGTGCTCGCTCC	330							
DB	8889	GGTCATGTTGACCATCGGAACACTACTTTAACCGAGCTCTAACCATGGCTCTCGCTTCC	8948							
QY	331	GTCCGGTGGCAGCGCCCTAAGAGTACGACGAGATCGACCGCGCCGCCCGGAGGAGCGGCC	390							
DB	8949	ATTGGTTCCAGCGTCGCTAAAGAGTACGAGAGATGACGTCGCGCGGAGAGAGCT	9008							
QY	391	CGCGGTATCACCATCAACACCGCCACCGTCGAGTACGACCGAGACCGCCGCACTACGCA	450							
DB	9009	CGTGGTATCACAATCAACACTGCTACTGTGAGTACGAGACTGAGAATCGTCACTACGCT	9068							
QY	451	CAGTCGACTGCCCGCGCCACCGCACTATGTCAAGAAATATGATCACCGCGCTCGCGCAG	510							
DB	9069	CACGTTGATTGCTCGTCACGCTGATTAGTTAAGATATGATTACCGAGCTGCACAG	9128							
QY	511	ATGACCGTGGCCATCTCGTGTATCCGGTGGCGACGGGCCCATGCCGAGACCAAGAG	570							
DB	9129	ATGACCGGAGCTATCTCGTTGTTCCGGCGCGATGCTCTATGCTCAGACTAAAGAG	9188							
QY	571	CACATCCTCTCGCCCAAGCAAGTCGGTGTCCCAAGATCGTTCTCTCTCAACAAGAG	630							
DB	9189	CATATCCTTTTGGCTAAGCAGGTTGGTGTCTCTGATATGTTGTCTTAAACAAGAG	9248							
QY	631	GACATGGTCACGACGAGGAGCTGCTCGAGCTCGAGCTCGAGCTCGCGAGCTGCTC	690							
DB	9249	GATCAAGTAGATGATGACAGTGTCTGAGAGCTGTTGAGCTTGGAGTTCGTGAGCTTCTC	9308							
QY	691	AGCACTACGAGTACGACGCGGACGACGATACCAATCGTCTGCTGCTCCGCCCTCAAGGCG	750							
DB	9309	TCGCTCTTATGAATTTAACGGTGATGATATTCGATATCTCTGTTCTCTCTTTTAGCC	9368							
QY	751	CTCGAGGCTCTCATGTCTAACCTCGCTTGAAGCGCGGCGAGCATGAGTGGTCTGACTAC	810							
DB	9369	GTTGAGACTCTTACTCAGAAATCCTAAGGTTAAGAGAGGTTGATAACAAATGGGTAGATAAG	9428							
QY	811	ATCTTCTCGTGGTGTATAAAGTGGATTCCTATATTCAGTCCCGCAGGAGGAGACTGAC	870							
DB	9429	ATTATGAATTTATGATGCTGTTGATGATTACATCCCTATCCCTCAGAGACAACTGAA	9488							
QY	871	CTCCCGTCTTCTGCTGTTGAAGATGCTTCTCCATCACCGGTGCTGATACAGTTGCC	930							
DB	9489	TTGCCATTTCTTTAGCTTGTAGGATGTTCTCTATCACCTGGACGCTGCTACGGTGGCT	9548							
QY	931	ACTGGCGGTATAGAGCGTGGCACCCTCAAGATTGGTGGACAGCTCGATATCGTCGGAATC	990							
DB	9549	ACAGGGCGTGTGAGAGAGGTTACGGTTAAGTAGGAGAGACTGTAGATTTAGTGGGTG	9608							
QY	991	CGGACACCCGGAACTGCAACGCTCACTGGTGTGTAGATGTTCCAGAAAGCACTGGATGAT	1050							
DB	9609	ACGACACTAGGAGTTACACTGTCACTGGGTGTAATGTTTCAGAGAGATTTCTGATGAG	9668							
QY	1051	GCCATGGCGGAGACAAATGTTGGGCTGCTGCTCCGTTGGTATGCAGAAAGATGACATGAA	1110							
DB	9669	GCTTTAGCTGGTGACAATGTAGGGTGTGCTTAGGGGTATTCAAAGGCTGATATTCAG	9728							
QY	1111	AGAGGATGGTGTGCAAAAGCCTGCTCTATCACACCGCACCAACCAAGTTTGGGCTGTT	1170							
DB	9729	AGAGGTATGGTTTTAGCTAAGCGGGATCGATTCTACTCCACATACCAAGTTTGAAGCAAT	9788							
QY	1171	GTGTATGTCTTAAAGGAAGAGGCTGCGCAGACTCACCTTCTCTCTCTGCTGTTACCGC	1230							
DB	9789	ATCTATGTGTTGAAGAAAGAGGAGGTTGAAGGCAATCTCCATTTCTTCAGGCTACAGG	9848							
QY	1231	CCACGTTCTTACATCGGCACTGATGTGACAGGAGGTGTGCTACGATATGATGAC	1290							
DB	9849	CCTCAGTTCTACATGAGGAGCTGATGTTTACGGGTAAAGTGACGAGAGATCATGACGAC	9908							

QY	1291	AAGGATGAGGAGCGAAGATGTCATGCGTGGTGACCGTATCAAAATGATTGTTACGCTC	1350
Db	9909	AAAGATGAAGAGTCGAAGATGTTATGCGCGTGATCGAGTGAAGATTGTTTGACGCTT	9968
QY	1351	ATCCAGCCTGTTGTTGTGACAGGATGATGAGGTTGCTATCCGTAGGCGTGAAGACC	1410
Db	9969	ATTGTGCGGGTGGCTTGTGAACAAGGATGAGGTTGCTATCAGAGAAGGAGGAAGACT	10028
QY	1411	GTTGTGCGCGTGTCATCAACAAATCATTCAGTAA	1446
Db	10029	GTTGTGCTGGAGTTATTGGGACGATCCTCGAATGA	10064
RESULT 15			
LOCUS	ATCHRIV52	198427 bp	DNA linear PLN 16-MAR-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 52.		
ACCESSION	AL161552		
VERSION	AL161552.2	GI:7268789	
KEYWORDS			
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 82704)		
AUTHORS	Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and		
	Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 67424 to 179368)		
AUTHORS	Terryn,N., Ardillies,W., Buysschaert,C., Dasseville,R., De Clerck,R.,		
	De Keyser,A., Neyt,P., Rouze,P., Van Den Daele,H., Villaroel,R.,		
	Gielen,J., Van Montagu,M., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	3 (bases 69292 to 69818)		
AUTHORS	Volckaert,G., Grymonprez,B., Voet,M., Robben,J., Mewes,H.W.,		
	Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	4 (bases 178975 to 187911)		
AUTHORS	Mewes,H.W., Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	5 (bases 187649 to 198427)		
AUTHORS	Pohl,T., Weizenegger,T., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	6 (bases 1 to 198427)		
AUTHORS	EU Arabidopsis sequencing project.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer		
	Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:		
	lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project		
	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge		
	Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,		
	E-mail: michael.bevan@bbsrc.ac.uk		
COMMENT	Information on performance of analysis and a more detailed		
	annotation of this entry and other sequences of chromosomes 3, 4		
	and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/		
	this fragment has an overlap with ATCHRIV51 at the 5' end and an		
	overlap with ATCHRIV53 at the 3' end.		
FEATURES	Location/Qualifiers		
Source	1..198427		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
gene	7547..8512		
	/gene="AT4g19950"		
join	(7547..8329,8375..8512)		
CDS	/gene="AT4g19950"		
	/note="similarity to p48 eggshell protein gene,		
	Schistosoma mansoni, PATCH:G45844		
	contains EST gb:AI994604.1, R30520, AI992766.1, T20423,		
	AA12864, H76323"		
	/codon_start=1		
	/product="putative protein"		
	/protein_id="CAB78995.1"		
	/db_xref="GI:7268790"		
	/translation="MDLAPEELQFLNKGILRESTSIPQYSLKTYFLTLTLPIPLSF		
	ALLAHSFTQPILAGIDTPQADQSQLOHEWTVLVFOFCYIIIFLAFSLSTAAVVF		
	TVASLYTGKPVSFSTMSAIPLVLRLEFITEVLSLLMLAYNTVFLIPVLIIVAVDL		
	QNVLIASFVSVFVFLVHVHTALWHLASVSVLEPIYGLAAMKSYELLGKTL		
	NACSMVFLVHVCFTAGVFGAVVVRGGDDYGITAFIRIVAGGFSLVYVYVCKSFHQEI		
	DKSALDHLGGVLGVEYVPLKNIQMFEN"		
	7547..8329		
	/gene="AT4g19950"		
	/number=1		
intron	8330..8374		
	/gene="AT4g19950"		
	/number=1		
exon	8375..8512		
	/gene="AT4g19950"		
	/number=2		
gene	11377..14567		
	/gene="AT4g19960"		
join	(11377..11526,11621..11846,11925..12170,12258..12332,		
CDS	12383..12643,12687..12887,12986..13288,13407..13929,		
	14024..14567)		
	/gene="AT4g19960"		
	/note="similarity to putative potassium transporter AtKT2p		
	& AtKT1p, Arabidopsis thaliana, Patch:G2384669 &		
	Patch:G2384671		
	contains EST gb:AI997834.1"		
	/codon_start=1		
	/product="potassium transporter-like protein"		
	/protein_id="CAB78996.1"		
	/db_xref="GI:7268791"		
	/translation="MAERVEASSPEGENTIEEREVGAMWELEKLDQPMDEANKLN		
	NYRKGLSMLMLLSFQSLGIVVDLGTPLVYVTFPDGDDSDSDFVIGALSLLI		
	YSLLIPLIKYVFIWCKANDNGOGTIAIYSLGRHAKVLIHQHRSDEDLTYSRT		
	VSAGSFAAKTKKWLGEKWRKRALLVVLGTCMGIDGILTFPAISYTGKIKYNNPK		
	MSGNHLFTLFTSDIVVVAIVLILIGLFSQHYGTQKVGWLFAPVILVILFICATG		
	MYNICKYDTSVLKAFSPITYIYLYFKRRGRDQWISLGILLISITAIKCYFLNIAGTEA		
	LYDIAYFPLLAIQLAFTFFPCLLAYCGAAYLVIRKHYQDAFYASIPDSVYWP		
	MFIVATGAAIVGSOATISGTYSIVKQVAHGCFPRKIVHTSKFLGQIYCPDINWIL		
	MLGCIANTAFSKQSOIGNAYGKMTTSKYKNYFSOWTAVVLMVLTLLMVLIMLL		
	VHCHWILVLIETLSFVSELSYSAVIEKIDEGWVPLIIAAILSLVMSVHYATVK		
	KYEFEMHRSKMSWILGSLGVRVPGIYVTELASGVPHIFSHFTNLPAIHVS		
	VYFVCYLPVTTPPEERFLVIRGPKTRFRFCRVARYGYKDLHKDDDDFENKLLTK		
	LSSTIRIEMPTSSNSTYSSTYSVNHQTQDSTVDLIHNNHNNHNNMDMFSSMDY		
	TVSTLDITVSAESLHNTYFSQDNTVEEETDELEFLTKCKESGVVHMGNTVVKART		
	GSWLPKTAIDYVYAFIAKICRANSVILHVPHTLLNVGQVYV"		
	11377..11526		
exon	/gene="AT4g19960"		
	/number=1		
intron	11527..11620		
	/gene="AT4g19960"		
	/number=1		
exon	11621..11846		
	/gene="AT4g19960"		
	/number=2		
intron	11847..11924		
	/gene="AT4g19960"		
	/number=2		
exon	11925..12170		
	/gene="AT4g19960"		
	/number=3		
intron	12171..12257		
	/gene="AT4g19960"		
	/number=3		
exon	12258..12332		
	/gene="AT4g19960"		
	/number=4		
intron	12333..12382		
	/gene="AT4g19960"		
	/number=4		
exon	12383..12643		

	Matches	892;	Conservative	0;	Mismatches	344;	Indels	0;	Gaps	0;
QY	211	GTGCGCGCGGAGGGGCGCAAGTTTCGAGCGCGCACCAACACACGTCACATAGCAGCACATC	270							
Db	187803	GTCCGCGCGCGTCTGCGAAAGTTTCGAGAGGAGAAAGCCTCATGTCAACATCGAACCATC	187862							
QY	271	GGCCATGTGACACGAGGAGACACTCTCACCGCGCGCTCACCATGTGTGCTCGCTCC	330							
Db	187863	GGTCATGTGACCATGGAAGAACTACTTAAACCGCAGCTTACCATGCTCTCGCTCC	187922							
QY	331	GTGCGTGGCAGCGCGCTTAAGAGTAGACGAGATGACGCGCGCGCGCGCGAGGCGCGC	390							
Db	187923	ATTGGTTCAGCGCTCGCTAAAGTAGTACGAGATTGACGCTGCGCGGAGGAGAGCT	187982							
QY	391	CGCGTATCACCATCAACACCGCGCGCTAGTAGAGACGAGACCGCGCGCGCGCGCGC	450							
Db	187983	CGTGTATCAACATCAACACGCGCTTGTGTGAGTACGAGACTGAGATCGTCACTACGCT	188042							
QY	451	CACGTGCACTGCCCGCGCGCGCTATGTCAAGAAATATGATCAACCGCGCGCTGCGCAG	510							
Db	188043	CACGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188102							
QY	511	ATGACGCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	570							
Db	188103	ATGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188162							
QY	571	CACATCTCTCTGCGCAACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	630							
Db	188163	CATATCTCTCTGCGCAACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188222							
QY	631	GACATGCTGCGACGCGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	690							
Db	188223	GATCAAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	750							
QY	691	AGCAACTAGGAGTAGGAGCGGCGAGCGAGTACCAATCGTCTGCTGCTGCTGCTGCT	750							
Db	188283	TCGCTTATGAATTTAAGCGGTGATGATATCCGATATATCTGCTGCTGCTGCTGCT	188342							
QY	751	CTGAGGCTCTCATGCTCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	810							
Db	188343	GTTGAGACTCTTACGTAGAACTCTTAAAGGAGTGTAAAGGAGTGTAAAGGAGTGT	188402							
QY	811	ATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	870							
Db	188403	ATTTATGAATTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188462							
QY	871	CTCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	930							
Db	188463	TTGCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188522							
QY	931	ACTGGCGGTATAGCGGTGGCGCGCTCAAGATTTGCTGCTGCTGCTGCTGCTGCT	990							
Db	188523	ACAGGCGGTGCTGAGAGAGGTACGCTTAAAGGTAGGAGAGCTGTAGATTTAGT	188582							
QY	991	CGGGACACCGGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1050							
Db	188583	AGGAGACTAGGAGTGTACACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188642							
QY	1051	GCCATGCGCGGAGACAAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1110							
Db	188643	GCTTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188702							
QY	1111	AGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1170							
Db	188703	AGAGGTATGCTTTTGTAGCTAAGCGCGGATCGATTTACTCCACATACCAAGTTT	188762							
QY	1171	GTCTATGCTGCTTAAAGGAGGAGGTGGCGCGCGCTGCTGCTGCTGCTGCTGCT	1230							
Db	188763	ATCTATGCTGCTTGAAGAAAGAGGAGGTGGAAGGCTTCTCCATTTCTTCCAGG	188822							
QY	1231	CCACAGTCTTACATCGCGCAACTGATGTGACAGGAGGTGTGCTGCTGCTGCTGCT	1290							
Db	188823	CCTCAGTCTTACATGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	188882							

Query Match 40.8%; Score 685.6; DB 8; Length 198427;
Best Local Similarity 72.2%; Pred. No. 3.7e-83;

Search completed: May 1, 2003, 09:37:37
Job time : 5544 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 15:58:03 ; Search time 480 Seconds

(without alignments)
7886.686 Million cell updates/sec

Title: US-09-810-764A-6

Perfect score: 1681

Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

Result No.	Score	Query Match %	Length	ID	Description
1	1405.4	83.6	1657	21 AAC46577	Zea mays DNA fragm
2	687.2	40.9	1651	21 AAC37985	Arabidopsis thalia
3	685.6	40.8	1651	21 AAC51055	Arabidopsis thalia
4	528.4	31.4	1194	23 AAS54260	Pseudomonas aerugi
5	527.8	31.4	1194	13 AAO20215	Sequence of tuf1 g
6	526.2	31.3	1194	13 AAO20219	Sequence of Srtufr
7	526.2	31.3	1194	13 AAO20218	Sequence of Srtufr
8	524.6	31.2	1194	13 AAO20220	Sequence of Srtufr
9	524	31.2	29379	23 AAS59510	Propionibacterium

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

10	521.6	31.0	1194	23 AAS54267	Pseudomonas aerugi
11	518.2	30.8	1194	13 AAO20221	Sequence of Srtufr
12	508.8	30.3	1185	21 AAO63694	E. coli proliferat
13	508.8	30.3	1185	22 AAH00707	Escherichia coli n
14	508.8	30.3	1185	23 AAS52628	E. coli DNA for ce
15	505.6	30.1	1230	23 AAS56229	Salmonella typhi D
16	503.6	30.0	1185	23 AAS52697	E. coli DNA for ce
17	498	29.6	1191	22 AAH00872	Bordetella pertuss
18	490	29.1	1194	13 AAO20216	Sequence of tuf2 g
19	488.4	29.1	4411529	22 AAI99682	Mycobacterium tube
20	484.4	28.8	1109	22 AAH01846	Klebsiella oxytoca
21	477.4	28.4	4403765	22 AAI99683	Mycobacterium tube
22	472	28.1	65632	21 AAA81502	N. meningitidis pa
23	472	28.1	349980	21 AAF21544	Neisseria meningit
24	472	28.1	1437668	21 AAA81490	N. meningitidis B
25	462	27.5	17782	21 AAA81530	N. meningitidis pa
26	460.6	27.4	1116	22 AAH01849	Obesumbacterium pr
27	455.8	27.1	1116	22 AAH01845	Buttiauxella agres
28	450.2	26.8	349980	22 AAH68525	C glutamicum codin
29	449.4	26.7	1191	22 AAH00677	Corynebacterium gl
30	446.6	26.6	1351	18 AAT92619	Elongation factor
31	446.4	26.6	1188	22 AAH68402	C glutamicum codin
32	443.4	26.4	1108	22 AAH01847	Plesiomonas shigel
33	436	25.9	1200	23 AAS53681	Helicobacter pylor
34	436	25.9	1448	19 AAY07964	Helicobacter pylor
35	433.4	25.8	49617	22 AAF28541	Genomic fragment #
36	432	25.7	1129	22 AAH01850	Budvicia aquatica
37	421.8	25.1	1112	22 AAH01533	Proteus mirabilis
38	420.4	25.0	1185	23 AAS53325	Haemophilus influe
39	420.4	25.0	1185	23 AAS53335	Haemophilus influe
40	420.4	25.0	1830121	17 AAT42063	Haemophilus influe
41	420.4	25.0	1830121	17 AAT42063	Haemophilus influe
42	418.4	24.9	1170	13 AAO20217	Sequence of tuf3 g
43	413.8	24.6	1057	24 ABK74018	Bacillus lichenifo
44	405.4	24.1	1107	22 AAH01848	Shewanella putrefa
45	402.8	24.0	1245	24 ABN93077	Staphylococcus epi

ALIGNMENTS

RESULT 1	
AAC46577	
ID AAC46577 standard; DNA; 1657 BP.	
XX	
AC AAC46577;	
XX	
DT 18-OCT-2000 (first entry)	
XX	
DE Zea mays DNA fragment SEQ ID NO: 50647.	
XX	
KW Hybridisation assay; genetic mapping; gene expression control;	
KW protein identification; signal transduction pathway; metabolic;	
KW pathway; promoter; termination sequence; corn; ss.	
XX	
OS Zea mays subsp. mays.	
XX	
PN Ep1033405-A2.	
XX	
PD 06-SEP-2000.	
XX	
PF 25-FEB-2000; 2000EP-0301439.	
XX	
PR 25-FEB-1999; 99US-0121825.	
PR 05-MAR-1999; 99US-0123180.	
PR 09-MAR-1999; 99US-0123548.	
PR 23-MAR-1999; 99US-0125788.	
PR 23-MAR-1999; 99US-0126264.	
PR 29-MAR-1999; 99US-0126785.	
PR 01-APR-1999; 99US-0127462.	
PR 06-APR-1999; 99US-0128234.	
PR 08-APR-1999; 99US-0128714.	
PR 16-APR-1999; 99US-0129845.	

PR	19-APR-1999;	99US-0130077.	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	30-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134257.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134218.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134219.	PR	27-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	18-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	19-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147303.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147416.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0147935.
PR	14-JUN-1999;	99US-0139119.	PR	10-AUG-1999;	99US-0148171.
PR	16-JUN-1999;	99US-0139452.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139453.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139454.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139455.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139456.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139457.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139463.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-0140353.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140354.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140695.	PR	01-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	99US-0140823.	PR	07-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	99US-0140991.	PR	10-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.

[illegible]

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19375.
 xx Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 XX EP1033405-A2.
 XX
 XX 06-SEP-2000.
 PD
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 03-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 24-MAY-1999; 99US-0135353.
 PR 25-MAY-1999; 99US-0135629.
 PR 27-MAY-1999; 99US-0136021.
 PR 28-MAY-1999; 99US-0136392.
 PR 01-JUN-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137502.
 PR 08-JUN-1999; 99US-0137724.
 PR 10-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.

```

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 40.9%; Score 687.2; DB 21; Length 1651;
Best Local Similarity 72.2%; Pred. No. 9.9e-135;
Matches 893; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY 211 GTGCGCGCGGAGGCGCAAGTTCGAGCGCACCAACACACAGCTCAACATAGGACCAATC 270
DB 261 GTCCGCGCGCGTCTGGAAGTTCGAGAGGAAGAGCCTCATGCAACATCGAACCATC 320
QY 271 GGCCATGTGACCAAGGAAGACCTCTCCACCGCGCGCTCACCATGCTGCTCGCCTCC 330
DB 321 GGTCTGTTGACCATGGAAGAACTACTTTAACCGGAGCTCTAACCATGGCTCTCGCTTC 380
QY 331 GTCCGTGGCAGCGGCTTAGAAGTAGCAGGAGATCGACCGCGCGCGCGCGGAGCGGCC 390
DB 381 ATTGGTTCCAGCGTCTGCTAAAAGTAGACGAGATTGACGCTCGCGCGGAGGAGAGCT 440
QY 391 CGCGGTATCAACATCAACCGCACCGTCTGAGTACGAGACCGAGACCGCGCCACTACGCA 450
DB 441 CGTGGTATCAATCAACATCTGCTACTGTTGAGTACGAGACTGAGATCGTCACTAGCT 500
QY 451 CAGCTGACTGCCCCGCCCGCGCTACTATGTCGAAGAATATGATCAACCGCGCTGCGCAG 510
DB 501 CACGTTGATGCTCTGCTGCTACGCTGATTAAGTAAAGATATGATTACCGGAGCTGCACAG 560

```

```

QY 511 ATGGACGGTGCATCTCTGCTATCTCGGTGCGAGCGGCCCATCCCGACGACCAAGAG 570
DB 561 ATGGACGGAGTATCTCTGTTTCCGCGCGCGATGCTCTATGCTCCTACACTAAGAG 620
QY 571 CACATCTCTCTGCGCAAGCAAGTGGTTCCTCCCAAGATGTTCTTCTTCAACAAGAG 630
DB 621 CATATCTTTTGGCTTAAGCAGGTGCTCTGATATGTTGTTCTTCTTAAACAAGAG 680
QY 631 GACATGGTCGACGAGGAGGCTGCTCGAGCTGCTCGAGCTCGAGGTCGAGGTCGCTC 690
DB 681 GATCAAGTAGATGTCAGAGTTCGAGGCTGTTGAGCTGTTGAGGTCGTTGAGGCTTCT 740
QY 691 AGCAACTAGAGTACGACGCGGACGACGATCAACAATCTGCTGGCTCGCGCTCAAGCG 750
DB 741 TCGCTTATGATTTAACGGTGATGATATTCGATATCTCTGGTCTCTCTTTAGCC 800
QY 751 CTCGAGGCTCTATGCTCAACCTGCTTGAAGCGCGGAGGATGATGAGGTCGACTAC 810
DB 801 GTTGAGACTCTTACTGAGAATCTTAAGGTAAAGAGAGGTGATAACAAATGGGTAGTAAG 860
QY 811 ATCTTCTGTTGGTGAATAAGTGGATTCCTATATTCAGTCCCGCAGGACGAGCTGAC 870
DB 861 ATTTATGAACCTTATGGATGCTGTTGATGATACATCTCTCCCTACGACGACAACTGAA 920
QY 871 CTCGCGTCTTCTGCTGCTTGAAGATGCTTCTCCATCACGCTGCTGGTACAGTTGCC 930
DB 921 TTGCCATCTCTGTTAGCTGTTGAGGATGTTCTCTATCACTGGAGCTGGTACGTTGCT 980
QY 931 ACTGCCGCTATAGCGGTGGCACCTCAAGATTTGTTGACACAGTGCATGCTCGGAATC 990
DB 981 ACAGGCGGTGTCGAGAGAGGTACGGTTAAGGTAGGAGAGCTGTAGATTTAGTGGCTTG 1040
QY 991 CGGGACACCCGAACTGCACGCTCACTGGTGTGAGATGTTCCCAAGAACCATGATGAT 1050
DB 1041 AGGGAGACTTAGGAGTTACACTGTCTCTGGGTTGAAATGTTTCAGAAGATCTTCTGATGAG 1100
QY 1051 GCATGGCGGAGACAAATGTTGGGCTGCTGCTCGCTGTTATGACAGAGATGACATTTGAA 1110
DB 1101 GCTTTAGCTGGTGACAAATGTTAGGTTGTTGCTTAGGGTATTCAAAAGGCTGATATTGAG 1160
QY 1111 AGAGCATGTTGCTGGCAAGCTTGGCTCTATCACACCGCACACCAAGTTTGGAGCTGTT 1170
DB 1161 AGAGGTATGTTGTTAGCTTAAGCGGGATGCTTACTTCCACATACCAAGTTTGAAGCAAT 1220
QY 1171 GTGATGCTTTAAGAAGGAGAGGTTGGCGCACACTCACCTTTCTTCTTCCCTGGTTACCGC 1230
DB 1221 ATCTATGTTTGAAGAAGAGAGGTTGGAAGGCTTCCCATTTCTTTCAGAGTTACAGG 1280
QY 1231 CCACAGTTCTACATCGGCAACTGATGTGACAGGAGTGTGACTACGATTATGATGAC 1290
DB 1281 CCTCAGTTCTACATGAGGAGCTGATGTTACGGTAAAGTACGAGACGATCATGACACG 1340
QY 1291 AAGGATGAGGCGCAAGATGTCATGCTTGGTACCGCTATCAAAATGATTGTTTCAGCTC 1350
DB 1341 AAGATGAAGAGTCAAGATGTTTATGCCCGGTGATCGAGTGAAGATGTTTCTGAGCTT 1400
QY 1351 ATCCAGCTGTTCTGTTGAGCAGGATGAGGTTTCTATCCGTTGAGGTTGTTGAAGAC 1410
DB 1401 ATTTGCCGCTGCTGTTGAACAAGGATGAGTTTGTCTATCAGAGAAGAGGAAAGACT 1460
QY 1411 GTTGTGCGGCTGTCATCAACAAATCATTTAGTAA 1446
DB 1461 GTTGTGCTGGAGTTATTGGGACGATCTCGAATGA 1496

```

```

RESULT 3
AAC51055
ID AAC51055 standard; DNA; 1651 BP.
XX
AC AAC51055;
XX
DT 18-OCT-2000 (first entry)
XX

```

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67105.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125786.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130445.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

DE Pseudomonas aeruginosa DNA for cellular proliferation protein #391.

XX Antisense; ds; prokaryotic cellular proliferation gene;

KW antibiotic; antibacterial; drug design.

XX Pseudomonas aeruginosa.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-PSDB; AAU36401.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 7897; 511pp; English.

PS The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence encodes an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1194 BP; 259 A; 361 G; 223 T; 0 other;

XX Query Match 31.4%; Score 528.4; DB 23; Length 1194;

XX Best Local Similarity 66.1%; Pred. No. 1.9e-101;

XX Matches 813; Conservative 0; Mismatches 381; Indels 36; Gaps 2;

XX 217 GCGGCGAGGCGAATTCGACGCGACCAACACCAACGCTCAACATAGGACCATCGGCGAT 276

DB 1 GTGGCTAAGGAAATTCGACGCTAACAACCGCAGCTCAACGCTCGGCGACCATCGGTCAC 60

XX 277 GTCGACCAAGGAAAGACACATCTCACCGCGCTCACCATGGTCTCGCTCGGTCGGT 336

DB 61 GTTGACCATGGACAGACCATCTGACCGCTGCTGACCAAGGCTGCTCCGATACCTCG 120

XX 337 GCGAGCGCGCTAAGAGTAGACAGATCGACGCGCCCGCCGAGGAGCGCGCCGCGGT 396

DB 121 GGTGTTCCGCTCGTCTTCGATCAGATCGACAACGCGCGGAGAAAGCCGCGGT 180

XX 397 ATCCATCAACACGCGGCGGCTCGAGTACGAGACGAGACCGCCCACTAGCAGACGTC 456

181 ATCACCATCAACACCTCGCACGTTGAATACGATTCGGCTGTTCTGCTCACTACGCCACGTT 240

457 GACTGCCCGCGCCACCGGACTATGTCAAGATATATCATCACGGGCTCGCGAGATGGAC 516

241 GACTGCCCGCGCTCACCGGACTACGTGAAGAACATATCACGGTGTCTCCCATAGTGGAC 300

517 GGTGCCATCTCGTCTATCCGGTGCAGCGGCCCATCGCGAGACCAAGAGCACATC 576

301 GCGCGATCTGTTGCTCGGCTGCCAGCGGCCCATGCCGAGACCCGCGAGCACATC 360

577 CTCCTGCCAAGCAAGTGTGTTCCCAAGATGTTCTTCTTCAACAAGAGACATG 636

361 CTGCTGTCCCGCGCAGGTAGCGCTTCCCTACATCGTCTGTTCTCTGAACAAGCCGACATG 420

637 GTCGAGCAGGAGGAGTGTCTCGAGCTCTCGAGCTCGAGTCCGAGGTCGCGAGTGTCTCAGCAAC 696

421 GTCGAGCAGCGGAGTGTCTGGAATCTGGTGGAGATGGAAGTTCGCGATCTGCTGAACAC 480

697 TAGAGTAGACGCGGACGACGATACCAATCGTCTGGCTCGCGCTCAAGCGGCTCGAG 756

481 TAGCACTTCCGGGCGGACGACACTCCGATCATCTCGGTTCCGCGCTGATGGCGCTGGA 540

757 GCTCTCATGCTCAACCTGCTTGAAGCGGCGGAGCAGTGAAGTGGTCTGCACTACATCTTC 816

541 G-----GCAAGGATGACAACGGCATCGGCGTAAGCGCCGCTGCAG 579

817 TCGTGTGTTGATAAATGATTCTTATTCAGTCCCGCAGAGGCGAGACTGACCTCCCG 876

580 AAGTGTGAGACCTGGACTCTTACATTCGCGACCGGTTGTCGCCATCGACCGACCG 639

877 TTCTTCTGCTGTTGAAGATGCTTCTCCATCACCGGCTCGGTGATGATGCTGCACTGCGC 936

640 TTCTGATCGGATGAAGACGTTCTCGATCTCGCGCGGTTACCGTGTGTAACCGT 699

937 CGTATAGCGGTGGACCGTCAAGATTTGTTGACACATCGATCGATTCGCGAATCCGGGAC 996

700 CGTGTAGCGGCGCATCATCAAGGTCAGGAAGAAAGTGAATCGTCGGCATCAAGGCG 759

997 ACCCGAATTCGACGCTGCTGTTGATGTTTCCAGAACACCATGATGATGCAATG 1056

760 ACCACCAAGACTACCTGACCGCGCTTGAATGTTCCGCAAGCTGCTCCAGCAAGTCTG 819

1057 GCGGAGACAATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116

820 GCTGTTGAGAACGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879

1117 ATGTTGCTGGCAAGCTGCTCTATCACACCGCACCAAGTTTTCAGGCTGTTGTGAT 1176

880 CAGGTACTGGCAAGCGGCGCACCATCAAGCGCGACACCAAGTTGAGTGCGAAGTGTAC 939

1177 GTGCTTAAAGAGGAGGTTGGCGCACACTCACCTTCTTCCCTGTTTACCGCCACAG 1236

940 GTGCTGTCCAAGGAAGAGTGTGCTGCTACACCCGCTTCTTCAAGGCTACCGTCCGCG 999

1237 TTCTACATCGGCAACTGATGTGACAGGAGTGTGACTACGATTATGATGATGATGATGAT 1296

1000 TTCTACTTCCGTACCGACCGATGTGACCG-----TAACCTGCGAACTGCGCG 1044

1297 GAGGAGCGCAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356

1045 GAAGGCGTAGAGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104

1357 CTGTTGCTTGTGAGCAGGTTATGAGTTTGTATTCGTTGAGGTTGTTGTTGTTGTTGTTG 1416

1105 CGATCGCATGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1164

1417 GCGGTTGCTCATCAACAAAATCATTTAGTAA 1446

1165 GCGGCGTGTGTTGCCAAGATCATCGAGTAA 1194

RESULT 5

XX DE Sequence of SrtuflR1 gene encoding elfamycin-resistant elingation
 XX factor EF-TuR Thr 378.
 KW Elfamycin resistant actinomycetes; antibiotic resistant;
 KW elongation factor; ss.
 XX OS Streptomyces ramocissimus.
 XX FH Key Location/Qualifiers
 FT CDS 4...1194
 FT /*tag= a
 XX EP466251-A.
 XX 15-JAN-1992.
 XX 02-JUL-1991; 91EP-0201702.
 XX 02-JUL-1991; 91EP-0201702.
 XX 10-JUL-1990; 90EP-0201851.
 XX (KONN) GIST-BROCADES NV.
 XX Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 PI Woudt LP;
 XX WPI: 1992-017874/03.
 XX P-PSDB; AAR20246.
 XX New protein conferring resistance to elfamycin - used to
 PT transform streptomycetes to resistant pheno-type
 XX Claim 8; Pages 13-15; 35pp; English.
 XX Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EF-TuR). The advantage of this change is that
 CC the limiting factor for the prodn. of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein
 CC resistant to an elfamycin, pref. mociimycin (Kirromycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.
 XX SQ Sequence 1194 BP; 239 A; 405 C; 373 G; 177 T; 0 other;

Query Match 31.3%; Score 526.2; DB 13; Length 1194;
 Best Local Similarity 67.2%; Pred. No. 5, 6e-101;
 Matches 832; Conservative 0; Mismatches 353; Indels 54; Gaps 4;

QY 217 GCGGCGAGGGCAAGTTGAGCGCCACCAACACACGTCACATAGGACCATCGGCCAT 276
 DB 1 GTGGCGAAGGGCGAAGTTGAGCGGACTAAGCGCACGTCACATGGGCACCATCGGTAC 60

QY 277 GTGCGACCGAGAAAGACCACTCTACCGCGCGCTCACCATTGGTGTCT-----GCCTCC 330
 DB 61 ATCGACCGAGTAAAGACGACCTCAGCGCGCCATACCAAGGTGCTGCACGACGCGTAC 120

QY 331 GTGCGTGGCAGCGCCCTAAGAAATGACGACGATGCGACGCGCGCGCGAGGCGGCC 390
 DB 121 CGGACCTGAAGCGGCCACCCCGTTGACACACATCGACAGGCTCTGTGAGGCGTAC 180

QY 391 CGCGGTATCACATCAACACCGCCACCGCTGAGTACGACGACCGCCACTAGCGCA 450
 DB 181 CGCGGTATCACATCTCCATCGGCACGCTGAGTACCGACCGAGCGGCGCTAGTAGCC 240

QY 451 CAGCTCGACTCGCCGCGCCAGCGGCTATGTCAAGAAATATGATCAGCGGCGTGGCCAG 510
 DB 241 CAGCTCGACTCGCCGCGTCAAGCGGCTATGATCAAGAAATATGATCAGCGGCGTGGCCAG 300

QY 511 ATGGAGCGTGCATCTCGTATCGGTGCGGCGGCGCCATGCGCGAGACCAAGAG 570
 DB 301 ATGGAGCGGCGCATCTCTGTGTCGCGGCCACCGACGCGCGGATGCGCGAGACCAAGAG 360

QY 571 CACATCTCTCTCGCCCAAGCAAGTCGGTGTTCCTCCAAAGATCGTTGTCTTCTCTCAACAAGAG 630
 DB 361 CACGTGCTCTCTCGCCCGCAGGTGCGGTTCGTATCATCGTGGTGGTCCCTGAACAAGGCC 420

QY 631 GACATGTCGACGACGAGGAGTGTCTCGAGCTCGTCGAGCTCGGAGTCCGCGAGCTGTC 690
 DB 421 GACATGTCGACGACGAGGAGTGTCTCGAGCTCGTTGAGCTCGTGGTTCGAGCTCGTCTC 480

QY 691 AGCAACTACGAGTACGACGCGGACGACGTACCAATCTGCTGCTGCTCGCCCTCAAGCG 750
 DB 481 TCCGAGTACGAGTTCGCGGCGCAGACCTGCGGTGCTCGCGGTCTCGCGCTGAAGCG 540

QY 751 CTCGAGGCTCTCATGGTCAACCCCTGCTTGAAGCGCGCGACGATGAGTGGGTGCTGACTAC 810
 DB 541 CTGGAG-----GGCGACGCTCAGTGCACGAGTCC 570

QY 811 ATCTTCTCGTTGTGATAAAGTGGATTCCTATATTCAGTCCCGCAGACGACACTGAC 870
 DB 571 GTCTCGACCTGATGAAGGCGCTGACGAGTCCATCCGCGGAGCGGAGCGGAGCTCGAC 630

QY 871 CTCGCGTTCTTCTCGCTGTTGAAGATGCTCTTCTCCATCAACCGTCTGCTGCTGCTGCTG 930
 DB 631 AAGCGTTCTCTCATGCGGATCGAGGAGCTCTTCAAGATCAACGCTCGCGGACGCTGCTC 690

QY 931 ACTGGCGTATAGAGCGTGGCACCGTCAAGATGCTGACAGTTCGATATGCTGCGGAATC 990
 DB 691 ACCGCGGTATCGAGCGTGTGCTCTGAAGGTCAACGAGACCGTTCGACATCATCGGCATC 750

QY 991 CGG---GACACCGCGAACTGCACGCTCAGTGTGCTGAGATGTTCCAGAAGACCATGGAT 1047
 DB 751 AAGACCGAGAAGACCAACACCGTCAACGCGATCGAGATGTTCCGAAGTGTCTCGAC 810

QY 1048 GATGCCATGGCGGAGACAATGTTGGCTGCTGCTGCTGATATGACAAGAGGATGACATT 1107
 DB 811 GAGGCGCAGGCGGTGAGAACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870

QY 1108 GAAAGAGCATGGTGTGGCAAGCGTCTATCATCACCGGACACCAAGTTGAGGCT 1167
 DB 871 GAGCGCGCAGGTTCATCATCAAGCGGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930

QY 1168 GTTCTGTATGTCTTAAAGAGAGGTTGGCGCGCACACTCAGCTTCTTCCCTGGTTAC 1227
 DB 931 CAGGCTTACATCTCTCCAAAGGAGGAGGTTGGCGCGCACACGCGTCTTTCACAACTAC 990

QY 1228 CGCCACAGTTCATACGCGGACAACTGATGTGACAGGAGTGTGACTACGATTATGAAT 1287
 DB 991 CGCGCGCAGTTCATCTCCGTCACGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1035

QY 1288 GACAAGGATGAGGAGGCGAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1347
 DB 1036 CACCTCCCGGAGGCGACCGAGATGTCATCGCGGCGGACAAACCGAGATGCGCGTCGAG 1095

QY 1348 CTGATCAGGCTGTGCTGTGAGCAGGATGATGAGTTTCTATCCGTGAGGTTGTAAG 1407
 DB 1096 CTGATCAGGCTGTGCTGTGAGCAGGAGGCTCAAGTTCACCATCCGTGAGGTTGCGCG 1155

QY 1408 ACCGTGTCGCGGTTGTCATCAACAAATCATTTGAGTAA 1446
 DB 1156 ACCGTGCGCGCGCGCGAGGTACCAAGATGCTCAAGTAA 1194

RESULT 7

AAQ20218

ID AAQ20218 standard; DNA; 1194 BP.

XX AAQ20218;

AC AAQ20218;

XX AAQ20218;

DT 15-APR-1992 (first entry)

XX Sequence of SrtuflR1 gene encoding elfamycin-resistant elingation
 DE factor EF-TuR Val 378.
 XX

KW Elfamycin resistant actinomycetes; antibiotic resistant;

KW elongation factor; ss.
 OS Streptomyces ramocissimus.
 XX
 XX Key Location/Qualifiers
 FT CDS 4..1194
 FT /*tag= a
 XX
 XX EP466251-A.
 XX
 XX 15-JAN-1992.
 XX
 XX 02-JUL-1991; 91EP-0201702.
 XX PF
 XX 02-JUL-1991; 91EP-0201702.
 XX PR
 XX 10-JUL-1990; 90EP-0201851.
 XX PR
 XX (KONN) GIST-BROCADES NV.
 XX PA
 XX
 XX Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 PI Woudt LP;
 XX
 XX WPI; 1992-017874/03.
 DR P-PSDB; AAR20245.
 DR
 XX New protein conferring resistance to elfamycin - used to
 PT transform streptomycetes to resistant pheno-type
 XX
 XX Claim 8; Pages 13-15; 35pp; English.
 XX
 CC Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EF-TuR). The advantage of this change is that
 CC the limiting factor for the prodn of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein
 CC resistant to an elfamycin, pref. mocimycin (Klirromycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.
 XX
 SQ Sequence 1194 BP; 238 A; 404 C; 374 G; 178 T; 0 other;
 XX
 Query Match 31.3%; Score 526.2; DB 13; Length 1194;
 Best Local Similarity 67.2%; Pred. No. 5.6e-101;
 Matches 832; Conservative 0; Mismatches 353; Indels 54; Gaps 4;
 QY 217 GCGCGAGGCGGCAAGTTCGAGCGACCAACACACAGTACGAGGCGCCATCGGCCAT 276
 DB 1 GTGCGAAGCGGAATTCGAGCGGACTAAGCCGACGTCACATGCGCACCATCGGTAC 60
 QY 277 GTGACACACGGAAGACCACTCTACCGCGCGCTCACCATGGTGTG-----CGCCTCC 330
 DB 61 ATGACACACGGAAGACGACCTTCAGGCGCGCATTTACCAAGGTGCTGCACGCGGTAC 120
 QY 331 GTGCGTGGCAGCGGCTTAGAAGTACGAGGATCGAGCGCGCCCGCCCGGAGCGCGCC 390
 DB 121 CGGACCTGACGAGCGGCGCCCGCTTCGACGACATGACAAAGGTCTCTGAGGAGCGTCAG 180
 QY 391 CGGCGTATACCACTAACCGCCGACCGCTGAGTACGAGACCGACCGCCACTACGCA 450
 DB 181 CGGCGTATACCACTCTCCATCGGCGACGCTGAGTACGAGCGGCGGTACGTACGCC 240
 QY 451 CAGCTGCACTGCCCCGCGCCACCGCGACTATGTCAGAAATATGATACCGCGCGCTGGCGAG 510
 DB 241 CAGCTGCACTGCCCCGCGGTACCGGCGACTACATCAAGAACAATGATCAGCGGTGCGCGCAG 300
 QY 511 ATGACCGGTGCCATCTCTGCTGATTCGGTGGCGGAGCGGCCCATGCGCGAGACCAAGAG 570
 DB 301 ATGACCGGCGCCATCTCTGTTGGTGGCGGCGACCGCGCCGATGCGCGAGACCAAGAG 360
 QY 571 CACATCTCTCTCGGCAAGCAAGTGGTGGTTCGCAAGATGCTTCTCTCTCAACAAGAG 630
 DB 361 CAGCTGCTCTGGCCCCCGGAGTGGGCTTCCGATACATGCTGGTGGTGGTGGGCGGCGG 420
 QY 631 GACATGCTCGACGACGAGGAGTCTCGAGCTGCTGAGCTCGTGGAGTTCGCGAGCTGCTC 690

DB 421 GACATGGTGGACGACGAGGAGATCATGGAGCTCGTGGAGTTCGAGGTCGAGCTCCTC 480
 QY 691 AGCACTACGAGTACGAGCGGACGACGTCACCAATCGCTGGTGGTGGTGGTGGTGGTGG 750
 DB 481 TCCGAGTACGAGTTCGCGGCGGAGACCTGCGGCTGCTCGGCTGCTCGGCTGCTCGGCT 540
 QY 751 CTCGAGGCTCTCATGGTCAACGCTGCTTGAAGCGCGCGACGATGATGGTGGTGGTGGT 810
 DB 541 CTGGAG-----GGCGACGCTCAGTGGACGACGCTCC 570
 QY 811 ATCTTCTCGTGGTGGTGAATAAGTGGATTCCTATATTCAGTCCCGCAGACGACACTGAC 870
 DB 571 GTCTCTCGACCTGATGAGGCGGTCGACGAGTCCATCCCGGAGCGCGGAGCGGAGCTCGAC 630
 QY 871 CTCGCGTCTTCTGCTGCTGTTGAAGATGCTTCTCCATCACCGGTGCTGGTGGTGGTGGT 930
 DB 631 AAGCGGTTCTCTCATGCCGATCGAGGACGCTTTCAGATCACCGGTGCGGCGGACGCTGTC 690
 QY 931 ACTGCGCGTATAGAGCGTGGCACCGCTCAAGATTTGGTACACAGTGGATGATTCGCGAATC 990
 DB 691 ACCGCGCGTATCGAGCGTGGTCTCTGAAGGTCAACGAGACCGCTCGACATCATCGGCATC 750
 QY 991 CGG---GACACCGGGAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047
 DB 751 AAGACCGGAGAGACCAACCGGTCACCGGTCACCGGTCACCGGTCACCGGTCACCGGTC 810
 QY 1048 GATGCCATGCGCGGAGACAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1107
 DB 811 GAGGCGCGGCGGCGGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
 QY 1108 GAAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1167
 DB 871 GAGCGCGCGGAGTCTATCATCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
 QY 1168 GTTGTGATGCTTGAAGAGAGAGGAGGTCGCGGCGGACACTTACCTTCTTCTTCTTCTT 1227
 DB 931 CAGGCGCTATCTCTCTCAAGAGAGGAGGTCGCGGCGGCGGCGGCTTCTTCAACAACAT 990
 QY 1228 CGCCACACATGCTTACATCGGGAACACTGATGTGACAGGAGTGTGACTACGATTTATGA 1287
 DB 991 CGCCCGCGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
 QY 1288 GACAAGGATGAGAGCGGAGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1347
 DB 1036 CACCTCCCGGAGGCGGACGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1095
 QY 1348 CTCATCCAGCTGTTGCTTGTGAGCAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1407
 DB 1096 CTGATCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1155
 QY 1408 ACCGTTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
 DB 1156 ACCGTCGCGCGCGGCGGAGTCAACGATCGTCAAGTAA 1194
 RESULT 8
 AAQ20220
 ID AAQ20220 standard; DNA; 1194 BP.
 XX
 AC AAQ20220;
 DT 15-APR-1992 (first entry)
 DE
 DE Sequence of SrtuR1 gene encoding elfamycin-resistant elongation
 DE factor EF-TuR Pro 378.
 XX
 KW Elfamycin resistant actinomycetes; antibiotic resistant;
 XX elongation factor; ss.
 OS Streptomyces ramocissimus.
 XX
 FH Key Location/Qualifiers

[illegible]

Pseudomonas aeruginosa DNA for cellular proliferation protein #390.
Antisense; ds; prokaryotic cellular proliferation gene;
antibiotic; antibacterial; drug design.
Pseudomonas aeruginosa.

OS Escherichia coli.
 XX WO200044906-A2.
 XX PD 03-AUG-2000.
 XX PF 27-JAN-2000; 2000WO-US02200.
 XX PR 27-JAN-1999; 99US-0117405.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX Zykind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI: 2000-514822/46.
 DR P-PSDB: AAB15890.
 XX Novel polynucleotides and polypeptides associated with microorganism
 PT proliferation, used to identify inhibitors of bacterial growth and
 PT proliferation, for use in antisense therapy.
 XX Claim 8; Page 115; 316pp; English.
 XX AAB65809 to AAB65889 and AAB66058 to AAB66138 represent nucleotide
 CC sequences derived from Escherichia coli which inhibit E. coli
 CC proliferation. AAB65890 to AAB66055 and AAB15886 to AAB16040 represent
 CC nucleotide and protein sequences associated with E. coli proliferation.
 CC AAB66056 and AAB66057 represent primers used for sequencing E. coli
 CC proliferation inhibiting nucleotide inserts in an example from the
 CC present invention. Methods from the present invention can be used to
 CC identify a proliferation- required gene in a microorganism, by contacting
 CC a microorganism with a proliferation-required gene activity inhibitory
 CC nucleic acid identified in another organism, and determining if
 CC inhibition occurs in the second microorganism. The nucleic acid sequences
 CC identified as being required for bacterial growth and proliferation, can
 CC be used for antisense therapy for killing bacteria.
 XX Sequence 1185 BP; 286 A; 306 C; 325 G; 268 T; 0 other;
 SQ

Query Match 30.3%; Score 508.8; DB 21; Length 1185;
 Best Local Similarity 65.3%; Pred. No. 2.5e-97;
 Matches 803; Conservative 0; Mismatches 382; Indels 45; Gaps 2;

Qy 217 GCGGGAGGGCAAGTTGAGCGCACCAACACACGCTCAACATAGGACACCATCGGCCAT 276
 Db 1 GTGCTAAAGAAATTTGAACGTFACAAACCGCACGCTTAACGTTGCTACTATCGGCCAC 60
 Qy 277 GTCGACCAACGAAAGACCACTCTCACCGCGCGCTCACCATGGTGTCTCGCTCGTGGT 336
 Db 61 GTTGACCAACGTAAGAACTACTCTGACCGCTGCAATCACCCGCTACTGGCTAAACCTAC 120
 Qy 337 GCGAGCGCGCTAAGAAAGTACGAGAGATCGACGCGCGCGCGCGGAGCGCGCGCGGT 396
 Db 121 GCGGTGCTGCTGTCGATTCGACCAATCGATACCGCGCGGAGAAAGCTCGTGGT 180
 Qy 397 ATCACCATCAACCGCCACCGCTCGAGTACGAGACCGCGCCGCTACTACGACACGTC 456
 Db 181 ATCACCATCAACACTTCTCAGTTGTAATACGACACCCCGACCGCTCACTACGACACGTA 240
 Qy 457 GACTTCCCCCGCCACCGCGCTACTGTCAAGAAATATGATACCGCGCGTGGCGAGATGGAC 516
 Db 241 GACTTCCCCCGCGCGCGCTACTGTGTAAGAAATATGATACCGCGCTGGCTGATGGAC 300
 Qy 517 GGTGCCATCTCTGCTGTATCCGTTGCGGACGCGCGCGCTACCGACCAAGACACATC 576
 Db 301 GCGCGATCTGTTAGTTGCTGCGATGACGCGCGCGCTGCGGACATCCTGAGCACATC 360
 Qy 577 CTCCTCCCAAGCAAGTCGTTGTTCCCAAGATCGTTTCTTCTCAACAAAGAGGACATG 636
 Db 361 CTGCTGGGTCGTACAGGTAGGCGTTCCGTCATCATCTGTTCTCTGAACAAATCGACATG 420
 Qy 637 GTCGACCAAGGAGTGTCTGAGCTCGTACGCTCGAGGTCCGCGAGTGTCTACGAAC 696

Db 421 GTTGATCAGCAGAGAGCTGCTGGAACCTGGTTGAAATGGAATTCGTGAACATTCGTCTCAG 480
 Qy 697 TACGAGTACGAGCGGACGACGCTACCAATCGTCGCTGCGCTCCGCCCTCAAGCGCTCGAG 756
 Db 481 TAGCACTTCCGGGCGACGACACTCCGATCGTTGCTGCTGCTGCTGAAAGCGCTGGA- 539
 Qy 757 GTTCTCATGGTCAACCCCTGCCTTGAAGCGCGGACGATGAGTGGGTCCGACTACATCTTC 816
 Db 540 -----AGCGAGCGAGATGGGAACGAAATCTCG 570
 Qy 817 TCGTTGGTTGAATAAGTGGATTCCTATATTCAGTCCCGCAGAGGACGACTGACCTCCG 876
 Db 571 GAACCTGGCTGGCTTCCTGGATTCTATATTCGGAACAGAGCGTGGATTGACAAGCG 630
 Qy 877 TTCTTCTGCTGCTTGAAGATGTCTTCTCATCACCGGTCGTTGAGTACAGTTCCTGCTG 936
 Db 631 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
 Qy 937 CGTATAGAGCGTGGCACCCTCAAGATTTGGTGACACAGTCGATATCGTCGGAATCCGGAC 996
 Db 691 CGTGTAGAAGCGGCTATCATCAAGTTGCTGAAGAGTTGAAATCGTTGTTCAAGAG 750
 Qy 997 ACCCGAATGCAAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
 Db 751 ACTCAGAAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
 Qy 1057 GCGGAGACAATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
 Db 811 GCTGCTGCAAGCTGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
 Qy 1117 ATGCTGCTGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176
 Db 871 CAGGTACTGGCTTAAGCGGCGCACCCTCAAGCGCGACACCAAGTTCGAATCTGAAGTAC 930
 Qy 1177 GTGCTTAAGAAGAGAGGCTGCGGACACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236
 Db 931 ATCTGTCCAAAGATGAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
 Qy 1237 TTCTACATCGGACAACTGATGTGACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1296
 Db 991 TTCTACTTCCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035
 Qy 1297 GAGGAGCGGAAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1356
 Db 1036 GAAGCGGTAGAGATGTTAATGCGGCGGACACATCAAAATGTTTACCTGATCCAC 1095
 Qy 1357 CTTGTTGCTGTTGAGCAGGATGAGTTTGTCTATCCGTGAGGGTGTGAAGACCGTTGGT 1416
 Db 1096 CGGATCGGATGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1155
 Qy 1417 GCGGTTGCTATCAACAAATCATTTGAGTAA 1446
 Db 1156 GCGGCGTTGTTGCTAAAGTCTGGGCTAA 1185

RESULT 13
 AAH00707
 ID AAH00707 standard; DNA; 1185 BP.
 XX AAH00707;
 XX 24-JUL-2001 (first entry)
 DE Escherichia coli nucleotide sequence SEQ ID NO:698.
 XX Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 XX vaccine; primer; ds.

RESULT 15

AAS56229 standard; DNA; 1230 BP.

XX AAS56229;

DT 13-FEB-2002 (first entry)

DE Salmonella typhi DNA for cellular proliferation protein #262.

XX Antisense; ds; prokaryotic cellular proliferation gene;

KW antibiotic; antibacterial; drug design.

XX Salmonella typhi.

OS Salmonella typhi.

XX WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.
DR P-PSDB; AAU38370.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 9866; 51lpp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1230 BP; 301 A; 323 C; 339 G; 266 T; 1 other;

XX Query Match 30.1%; Score 505.6; DB 23; Length 1230;

XX Best Local Similarity 65.1%; Pred. No. 1.2e-96;

XX Matches 801; Conservative 0; Mismatches 385; Indels 45; Gaps 2;

XX 216 CGCGGCGAGGGGCAAGTTCGAGCGCCACCAACACAGGTCAACATGAGGACCATCGGCCA 275

DB 45 CGTGTCTAAAGAAAAATTTGAAGGTACAAAACCGCGAGTTAAGCTGGCGACCATCGGCCA 104

QY 276 TGTGACACGGAAGAACCACTCTCACCGCGCGCTCACCATGTGCTCGCTCGCG 335
DB 105 CGTTGACACGGAAGAACCACTCTCACCGCGCGCTCACCATGTGCTCGCTCGCG 164
QY 336 TGGCAGCGCGCTAAGAAATAGCAGCAGATCGACCGCGCGCGCGCGCGCGCGCG 395
DB 165 CGCGGTGCTGCTCGTTCGATTCGACAGATCGATAACGTCGGAAGAAAAGCGGTGG 224
QY 396 TATCACCATCAACACCGCGCGCTGAGTACGAGACCGACCGCGCGCGCGCGCG 455
DB 225 TATCACCATCAACACTTCTACGTTGAATAGATACCGCGCGCGCGCGCGCGCG 284
QY 456 CGACTGCCCGCGCGCGCGCGCTATGTCAAGAAATATGATCAGCGCGCGCGCGCG 515
DB 285 AGACTGCCCGCGCGCGCGCGCTATGTAAACAAATATGATCAGCGCGCGCGCG 344
QY 516 CGGTGCGATCTCTGCTGCTATCCGCTGCGCGCGCGCGCGCGCGCGCGCGCG 575
DB 345 TGGCGCGATCTCTGCTGCTGCTGCGACTGACCGCGCGCGCGCGCGCGCGCG 404
QY 576 CCTCTCGCCAAAGCAAGTCTGCTTCCCAAGATCTGCTTCTCTCAACAAGAGCAT 635
DB 405 CCTGCTGGTCTGCTAGGTAGCGCTTCCGTACATCATCTGCTTCTGACAAATGCGACAT 464
QY 636 GGTGACGACGAGGAGCTGCTCGAGCTGCTCGAGCTCGAGGTCCGCGAGCTGCTCAGCAA 695
DB 465 GGTGATGACGAAGAGCTGCTGGAATGTTGAAATGGAAGTTCGTGAATCTCTCTCA 524
QY 696 CTACGAGTACGACGCGCGCGCGCTACCAATCTGCTGCTGCTGCTGCTGCTGCTG 755
DB 525 GTACGACTTCTCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 584
QY 756 GCCTCTCATGGTCAACCTGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 815
DB 585 -----AGCG 614
QY 816 CTCGTTGTTGATAAAGTGGATTCTTATATCCAGTCCCGCGAGAGCGCGCGCGCG 875
DB 615 CGAAGTGGTGGCTTCTCTGATTTTACATCCCGGAACAGAGCGGTGCGGATGACAAGCC 674
QY 876 GTTCTGCTGCTGTTGAAGATGCTTCTTCCATACCGCGTCTGCTGCTGCTGCTGCTG 935
DB 675 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
QY 936 CCGTATAGAGCGTGCACCGCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 995
DB 735 TCGTGTAGAACCGCGGTATCATCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCG 794
QY 996 CACCGGAACTCAGCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
DB 795 GACTCAGAGTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
QY 1056 GCGCGGAGCAATGTTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1115
DB 855 TCGTGTGAGAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 914
QY 1116 CATGCTGCTGCGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
DB 915 TCAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
QY 1176 TGTGCTTAAGAAAGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
DB 975 CATCTGCTCAAGATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1034
QY 1236 GTTCTACATGCGGACAACTGATGTGACAGGAGTGTGACTACGATATGAATGACAAGA 1295
DB 1035 GTTCTACTTCCGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079
QY 1296 TGAGGAGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1355
DB 1080 GGAAGCGGTAGAGTGGTAAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1139

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:06:31 ; Search time 2893 Seconds
(without alignments)
9410.513 Million cell updates/sec

Title: US-09-810-764A-6
Perfect score: 1681
Sequence: 1 attcccaataatcccccacc.....gttaaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_esthu:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1054	62.7	1193	11	AY105509
2	569.8	33.9	831	12	BE705560
3	563.4	33.5	624	9	AI665388
4	537.6	32.0	681	14	BQ245968
5	523.4	31.1	736	10	AV943931
6	521.2	31.0	761	12	BE705674

C	7	520	30.9	728	13	BJ308303	BJ308303
C	8	515.4	30.7	900	12	BG301275	BG301275
C	9	509.6	30.3	685	13	BJ263237	BJ263237
C	10	507.4	30.2	658	13	BJ302467	BJ302467
C	11	501.4	29.8	678	13	BJ959687	BJ959687
C	12	500.6	29.8	644	10	AV938814	AV938814
C	13	498.4	29.6	670	13	BJ255324	BJ255324
C	14	497.6	29.6	632	12	BG908358	TArl1167E
C	15	493.2	29.3	558	13	BJ926556	BJ926556
C	16	493	29.3	682	13	BJ320206	BJ320206
C	17	491.8	29.3	623	11	AY107645	AY107645
C	18	488.6	29.1	574	10	AW672043	AW672043
C	19	487.2	29.0	636	13	BJ217502	BJ217502
C	20	477.8	28.4	606	13	BJ257818	BJ257818
C	21	477.2	28.4	826	12	BF262018	BF262018
C	22	475.4	28.3	624	13	BJ249074	BJ249074
C	23	472.6	28.1	650	9	AL507013	AL507013
C	24	471.8	28.1	608	10	AV941424	AV941424
C	25	471.2	28.0	907	10	BE216372	BE216372
C	26	469.4	27.9	514	12	BG410506	BG410506
C	27	462.8	27.5	639	13	BJ118749	BJ118749
C	28	461	27.4	574	13	BJ208789	BJ208789
C	29	458.2	27.3	648	13	BJ946293	BJ946293
C	30	455.4	27.1	639	14	BQ789395	BQ789395
C	31	454.8	27.1	636	9	AJ431953	AJ431953
C	32	447	26.6	637	9	AL500399	AL500399
C	33	441.6	26.3	601	13	BJ229928	BJ229928
C	34	439.8	26.2	532	10	AV922915	AV922915
C	35	434.6	25.9	595	10	AV946486	AV946486
C	36	432.2	25.7	473	12	BG410505	BG410505
C	37	431.8	25.7	534	12	BF291714	BF291714
C	38	430.2	25.6	467	12	BG349916	BG349916
C	39	429.6	25.6	549	9	AL822168	AL822168
C	40	429.6	25.6	551	12	BG263489	BG263489
C	41	423	25.2	511	10	BE470937	BE470937
C	42	422.6	25.1	582	10	AV945813	AV945813
C	43	421	25.0	580	12	BG906350	BG906350
C	44	420.6	25.0	1815	11	AY104241	AY104241
C	45	418.2	24.9	712	13	BI941934	BI941934

ALIGNMENTS

RESULT 1
AY105509
LOCUS AY105509 1193 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PC0100031 mRNA sequence.
ACCESSION AY105509
VERSION AY105509.1 GI:21208587
KEYWORDS HTC
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1193)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
AUTHORS
Maize Mapping Project/Dupont Consensus Sequences for Design of
TITLE
Overgo Probes
JOURNAL
REFERENCE 2 (bases 1 to 1193)
Unpublished (2002)
AUTHORS
Coe,E.C.
TITLE
Direct Submission
JOURNAL
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA

FEATURES
source

Location/Qualifiers
1..1193
/organism="Zea mays"
/db_xref="MaizeDB:635741"
/db_xref="taxon:4577"
/clone="PC0100031"

```

/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT      283 a   290 c   328 g   291 t   1 others
ORIGIN

Query Match      62.7%; Score 1054; DB 11; Length 1193;
Best Local Similarity 93.4%; Pred. No. 4.4e-200;
Matches 1112; Conservative 0; Mismatches 76; Indels 3; Gaps 1;

QY 491 TGATCACCGCGCTGCGCAGATGCGAGCGGTCCATCTCTGCTATCCGCTGCGGCGGCG 550
Db 1 TGATCACCGCGCTGCGCAGATGCGAGCGGTCCATCTCTGCTATCCGCTGCGGCGGCG 60
QY 551 CCATGCGCGAGACCAAGAGACATCTCTCTCGCCCAAGCAAGTCTGCTCCCAAGATCG 610
Db 61 CCATGCGCGAGACCAAGAGACATCTCTCTCGCCCAAGCAAGTCTGCTCCCAAGATCG 120
QY 611 TTGCTCTTCTCAACAAGAGGACATGCTCGACGAGGAGGAGTCTGCTGAGCTGCTGAGC 670
Db 121 TTGCTCTTCTCAACAAGAGGACATGCTCGACGAGGAGGAGTCTGCTGAGCTGCTGAGC 180
QY 671 TCGAGTCCGCGAGCTGCTCAGCAACTAGAGTACGAGGAGGAGGAGGAGTCTGCTGAGC 730
Db 181 TGAAGTCCGCGAGCTGCTCAGCAACTAGAGTACGAGGAGGAGGAGGAGTCTGCTGAGC 240
QY 731 CTGGCTCCCGCTCAAGGCGCTCGAGGCTCTCATGTTCAAGCTGCTGAGGCGGCGG 790
Db 241 CTGGCTCCCGCTCAAGGCGCTCGAGGCGCTCTCATGTTCAAGCTGCTGAGGCGGCGG 300
QY 791 ACGATGAGTGGTGCATCATCTCTCTGCTGTTGATTAAGTGGATTCCTATATTCAG 850
Db 301 ACGATGAGTGGTGCATCATCTCTCTGCTGTTGATTAAGTGGATTCCTATATTCAG 360
QY 851 TCCCGCAGAGGAGTGCATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 910
Db 361 TCCCGCAGAGGAGTGCATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 911 CGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
Db 421 CGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 971 CAGTGCATATCTCGGATCCGGACACCCGGAACCTGACGCTGCTGCTGCTGCTGCTGCTGCT 1030
Db 481 CAGTGCATATCTCGGATCCGGACACCCGGAACCTGACGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 1031 TCCAGAAGACCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
Db 541 TCCAGAAGACCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1091 TCCAGAAGACCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
Db 601 TCCAGAAGACCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1151 ACACCAAGTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
Db 661 ACACCAAGTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1211 CTTTCTTCCCTGGTTACCGCCAGTTCATCATGCGGACCACTGATGCTGACAGGAGTG 1270
Db 721 CTTTCTTCCCTGGTTACCGCCAGTTCATCATGCGGACCACTGATGCTGACAGGAGTG 780
QY 1271 TGACTACGATTTGAATGACAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1330
Db 781 TGACTACGATTTGAATGACAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 840
QY 1331 TCAAAATGATTTGACGCTATCAGCCTGTTGCTGAGAGGAGGATGAGGTTGCTA 1390
Db 1391 TCAAAATGATTTGACGCTATCAGCCTGTTGCTGAGAGGAGGATGAGGTTGCTA

```

```

Db 841 TCAGATGGTGTTCAGCTCATCCAGCCTGTTGCTGTGACAGGGTATGAGGTTGCTA 900
QY 1391 TCCGAGGGGTGAAGACCGTTCGTCGCGGTGTCATCAACAAATCATTTGAGTAACTG 1450
Db 901 TCCGAGGGGTGAAGACCGTTCGTCGCGGTGTCATCAACAAATCATTTGAGTAACTG 960
QY 1451 GATATAACATATCCACCATGAGAAATTTCTTCTGTTACTCAAGCGACATGCTCCGAGT 1510
Db 961 GATATAACATATCCACCATGAGAAATTTCTTCTGTTACTCAAGCGAAATGCTCTGTAGT 1020
QY 1511 TCTTATTATGTTGAGTTCAGTTTGGGTTGCTCATGTCGCAATGTTAGTATGACACTTTT 1570
Db 1021 TCTTATTATGTTGAGTTCAGTTTGGGTTGCTCATGTCGCAATGTTAGTATGACAC 1077
QY 1571 TTTGTCAGTGAATTTGATTAATTTATGACATTCACGACAAAGATTCACATATGTTG 1630
Db 1078 TTTGTCAGTGAATTTGATTAATTTATGACATTCACGACAAAGATTCACATATGTTG 1137
QY 1631 CAACCTCATTTGGCTAAGAGTGCATCTACTGTTTAAAAAATAAAAAA 1681
Db 1138 CAACCTCATTTGGCTAAGAGTGCATCTACTGTTTAAAAAATAAAAAA 1188

RESULT 2
BE705560/c
LOCUS
DEFINITION
  BE705560 831 bp mRNA linear EST 12-SEP-2000
  Secale cereale cDNA clone Sc01_03a09, mRNA sequence.
ACCESSION
  BE705560
VERSION
  BE705560.1 GI:10093825
KEYWORDS
  EST.
SOURCE
  rye.
ORGANISM
  Secale cereale
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Secale.
  J.I., Ouellet,T., Piche,C., Couroux,P., De Moors,A., Harris,L.J., Hattori
  ,J.I., Ouellet,T., Robert,L.S., Spratt,D. and Tinker,N.A.
  Expressed Sequence Tags from Cold-Stressed Winter Rye Seedlings
  Unpublished (2000)
  Contact: Singh,J.A.
  Eastern Cereal and Oilseed Research Centre
  Agriculture and Agri-food Canada
  KW Neathly Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
  0C6, Canada
  Tel: (613) 759-1662
  Fax: (613) 759-1701
  Email: singhja@em.agr.ca.
  Location/Qualifiers
    1. 831
    /organism="Secale cereale"
    /cultivar="Puma (winter rye)"
    /db_xref="taxon:4550"
    /clone_lib="Sc01_03a09"
    /tissue_type="leaf, crown"
    /dev_stage="seedling three-leaf stage"
    /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
    Site_2: Xho I; Sampled three-leaf seedlings treated for
    one week at 20C, 12 hrs light/day. Library made with
    Stratagene UNIZAP XR Kit/GigaPack III Gold Kit. Lambda
    library is amplified, then mass excised in SOLR cells."
BASE COUNT      210 a   243 c   187 g   190 t   1 others
ORIGIN

Query Match      33.9%; Score 569.8; DB 12; Length 831;
Best Local Similarity 83.9%; Pred. No. 1.4e-103;
Matches 692; Conservative 1; Mismatches 123; Indels 9; Gaps 4;

QY 750 GCTGAGGCTCTCATGGTCAACCCCTGCTTGAAGCGCGGACGATGAGT-GGGTCGACT 808
Db 808 GCTGAGGCTCTCATGGTCAACCCCTGCTTGAAGCGCGGACGATGAGT-GGGTCGACT

```


Db 831 GCTCAGSCCCTCATGGCCACCCCTGGCCTCAAGCGTGGGATAACGAGTGGGGTGATG 772
 QY 809 ACATCTTCCTGCTGGTTGATGAAGTGGATTCCTATATTCAGTCCCGCAGAGGACAGTG 868
 Db 771 GCATCTTCCTGCTGGTTGATGACTCCGTCGATACCCATCCCTGCTCCGAGAGGACAGCG 712
 QY 869 ACCTCCCGTCTTCTGCTGCTGTTGAAGATGCTCTTCCATCACCGGTGCTGTACAGTTG 928
 Db 711 ACCTACCTCTTCTGCTGCTGTTGAGATGCTCTTCCATCACTGGCGGTGATAGTTG 652
 QY 929 CCACTGGCGTATAGAGCGTGGACCGTCAAGATTGGTGACACAGTCGATATCGTCGAA 988
 Db 651 CCACTGGCGTATAGAGCGTGGACCGTCAAGTTGG-GAACCACTCGACCTCGTCGGCA 593
 QY 989 TCCGGACACCCCGAAGTCCACCGTCACTGGTTGGATGTTCCAGAGACCATGGATG 1048
 Db 592 TCAGGAGACCCCGAAGTCCACCGTCACTGGTTGGATGTTCCAGAGACCATGGATG 533
 QY 1049 ATGCCATGGCGGAGACAATGTTGGGCTGCTGCTCGTGTATGCGAAGGATGACATTG 1108
 Db 532 ATGCTATTGCTGGGACATGTTGGACTGCTGCTCGTGTATGCGAAGGATGATG 473
 QY 1109 AAGAGCGATGCTGGCGAAGCCCTGGCTCTATCACACCGCACACCAAGTTTGAGGCTG 1168
 Db 472 AGAGAGCGATGCTGGCGAAGCCCGTTCATCACCGCACACCAAGTTTGAGGCTG 413
 QY 1169 TTGCTGATGCTTAAAGAGGAGAGGTTGGCGGACACTCACTTCTCCCTGGTTACC 1228
 Db 412 TTGCTTATGCTCAAGAGGAGGAGGTTGGCGGACACTTCCCTGGTTACC 353
 QY 1229 GCCACAGTCTACATCGCGAAGTGTGACAGGAGTGTGACTACGATTATGAATG 1288
 Db 352 GTCGCAATCTACATGAGGACTACTGATGTCAGGGGATGTGACAAACATATGAAG 293
 QY 1289 ACAAGGATGAGGAGGAGAGATGTGATGCTGCTGAGCGTATCAAAATGATTGTTACG 1348
 Db 292 ACAAGGATGAGGAGGAGAGATGTGATGCTGCTGAGCGTATCAAAATGATTGTTACG 233
 QY 1349 TCATCAGCCTGTTGCTGAGCAGGATGAGGTTGCTATCCGTGAGGTTGTTAAG 1408
 Db 232 TCATCAGCAGTGGCTGTGAGCAGGAGTGAAGTTGCCATCCGTGAAGTGGGAAGA 173
 QY 1409 CCGTGTGCGGCTGTCATCAACAAATCATTCAGTAACTGGATATACATATCCACCA 1468
 Db 172 CAGTGGTGCAGTGTGTCATCAATATATCATTCAGTAAAGTGGATGGAGATCCACCG 113
 QY 1469 TGAGAAATTTCCCTGTTTACTC-AAAGCGACATGCTCCGTAGTGTGTTATATGTTGAG 1527
 Db 112 TGAGAAATTTCCCTGTTTACTC-TTTTTCGAAATGCTATGATGTTGTTATATGCG- 58
 QY 1528 TTTTAGGGTGTGCTCATGTCGAATTTGATGATGACACTTTTTTTT 1572
 Db 57 -TTTAGGAGGCTCTGTGAAATTTGATGATGGAACATTTCT 14

RESULT 3
 AI665388/c 624 bp mRNA linear EST 02-FEB-2000
 LOCUS 605010F02.x1 605 - Endosperm cDNA library from Schmidt lab Zea mays
 DEFINITION cDNA, mRNA sequence.
 ACCESSION AI665388
 VERSION AI665388.2 GI:4804516
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 624)
 Walbot,V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 JOURNAL Unpublished (1999)

COMMENT On May 11, 1999 this sequence version replaced gi:4776385.
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 605010 row: F column: 02.
 FEATURES
 source
 1. 624
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="605 - Endosperm cDNA library from Schmidt lab"
 /tissue_type="nucellar, embryo, and endosperm"
 /dev_stage="10-14 days post-pollination"
 /lab_host="DH5(alpha)"
 /note="Organ: Kernel; Vector: pAD-GAL4-2'; Site:1; EcoRI;
 Site_2: XhoI; Kernel endosperm cDNA library from Schmidt
 lab"

BASE COUNT 169 a 179 c 123 g 152 t 1 others
 ORIGIN
 Query Match 33.5%; Score 563.4; DB 9; Length 624;
 Best Local Similarity 94.8%; Pred. No. 2.8e-102;
 Matches 582; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 956 TCAAGATTGGTGACACAGTCGATATCGTCGGAATCCGGGACACCCGGAACTGCGACGTCA 1015
 Db 624 TCAAGATTGGTGACACAGTCGATATCGTCGGAATCCGGGACACCCGGAACTGCGACGTCA 565
 QY 1016 CTGGTGTGAGATGTTCCAGAGACCATGGATGATGCCATGCCGACACAAATGTTGGCG 1075
 Db 564 CTGGGTTGAGATGTTCCAGAGACCATGGAGGATGCCATGGCTGGAGACAAATGTCGGAC 505
 QY 1076 TGCTGCTCCGTGGTATGCAAGAGGATGACATTGAAAGAGGATGCTGTCGCAAGCCTG 1135
 Db 504 TGCTGCTCCGTGGTATGCAAGAGGATGACATTGAGAGGATGCTGTCGCAAGCCTG 445
 QY 1136 GCTCTATCACACCCGACACCAAGTTGAGGCTGTTGCTATGCTTAAAGAGGAAGAGG 1195
 Db 444 GCTCTATCACACCCGACACCAAGTTGAGGCTGTTGCTATGCTTAAAGAGGAAGAGG 385
 QY 1196 GTGGCCGACACTCACTTTCTTCCCTGCTACCGCCACAGTCTACATGCGGACAACTG 1255
 Db 384 GTGGCCGACACTCACTTTCTTCCCTGCTACCGCCACAGTCTACATGAGGACAACTG 325
 QY 1256 ATGTGACAGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGGAAGATGCA 1315
 Db 324 ATGTGACAGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGGAAGATGCA 265
 QY 1316 TGCTGTGACCGTATCAAAATGATGTTGCTAGCTCATCCAGCCTGTTGCTGTGAGCAGG 1375
 Db 264 TGCTGTGACCGTATCAAAATGATGTTGCTAGCTCATCCAGCCTGTTGCTGTGAGCAGG 205
 QY 1376 GTATGAGTGTCTATCCGTGAGGCTGTTAAGACCTGTTGGTCCGCTGCTATCAACAAA 1435
 Db 204 GTATGAGTGTCTATCCGTGAGGCTGTTAAGACCTGTTGGTCCGCTGCTATCAACAAA 145
 QY 1436 TCATTGAGTAAACTGGATATACATATCCACCAAGAAATTTCCCTGTTTACTCAAAAGC 1495
 Db 144 TCATTGAGTAAACTGGATATATCATATCCACCAAGAAATTTCCCTGTTTACTCAAAAGC 85
 QY 1496 GACATGCTCCGTAGTGTGTTATGTTGAGGTTTGTAGGTTGCTCATGTGCAATGTA 1555
 Db 84 GAAATGCTCTGTAGTGTGTTATGTTGAGTGTGTTAGGTTTGTCTTGTGAAATGTA 25
 QY 1556 GTATGACACTTTTT 1569
 Db 24 GTATGACACTTTTT 11

RESULT 4
 BQ245968
 LOCUS
 DEFINITION
 BQ245968
 TAE15017G12R TAE15 Triticum aestivum CDNA clone TAE15017G12R, mRNA
 sequence.
 BQ245968
 BQ245968.1 GI:20411844
 EST.
 SOURCE
 ORGANISM
 Triticum aestivum
 bread wheat.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Triticum.
 1 (bases 1 to 681)
 Cloutier, S.
 Wheat functional genomics - Glenlea developing seeds cDNA libraries
 Unpublished (2002)
 Contact: Dr. Sylvie Cloutier
 Cereal Research Centre, Agriculture and Agri-food Canada
 195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
 Tel: (204) 983-2340
 Fax: (204) 983-4604
 Email: scloutier@em.agr.ca
 was cloned directionally, not all sequences generated with reverse
 primer were from the 5' end (same with forward primer and 3' end).
 Average insert size is >1.4 kb
 Plate: 017 row: G column: 12
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. 681
 /organism="Triticum aestivum"
 /cultivar="Glenlea"
 /db_xref="taxon:4565"
 /clone="TAE15017G12R"
 /clone_lib="TAE15"
 /tissue_type="developing seeds"
 /dev_stage="15 days after anthesis"
 /lab_host="E. coli DH10B"
 note=Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
 Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds
 of cultivar Glenlea 15 days post-anthesis"
 149 a 165 c 207 g 160 t

FEATURES

source

Query Match 32.0%; Score 537.6; DB 14; Length 681;
 Best Local Similarity 86.9%; Pred. No. 3.8e-97;
 Matches 591; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 QY 768 CAACCTGCTTGAAGCGCGGACGATGAGTGGTGGCTACATCTCTCTGTTGTTGA 827
 DB 1 CACCCCTGGCTCAAGCTGGGGATAACGAGTGGTGGTGGATCTCTCTTGGATTGA 60
 QY 828 TAAAGTGATTCCTATATCCAGTCCCGCAGAGGACAGTACCTCCGCTTTCTGTCGC 887
 DB 61 CTCCTGGATACCCACATCTCTGTCGCGCAGAGGACAGTACCTACCTCTTCTGTCGC 120
 QY 888 TGTTCAGATGCTCTCTCCATCAGCGTGGTGGTACAGTCCCTGCGCTATAGAGCG 947
 DB 121 TGTTCAGATGCTCTCTCCATCAGCTGGTGGTGGTACAGTTCCTGCTGCTATCGAGCG 180
 QY 948 TGGCACCCTCAAGATTGTTGTGACAGTGGATATGCTCGGAATCCGGACACCCGGAAGT 1007
 DB 181 TGGCACCCTCAAGTGGGGACCCAGTCCGCTGGCATAGGAGACTCCGAATGC 240
 QY 1008 CACGCTCACTGGTGTGAGATGTTCCAGAGACCATGGATGATGCCATGGCGGAGACAA 1067
 DB 241 CACGCTCACTGGTGTGAGATGTTCCAGAGACCATGGATGATGCCATGCTGGGACAA 300
 QY 1068 TGTTCGGCTGCTCTCTGCTGATGTCAGAGGATGACATTTGAAGAGCATGGTCTGGC 1127
 DB 301 TGTTCGGCTGCTCTCTGCTGATGTCAGAGGAGATGTCAGAGGAGCATGGTCTGGC 360
 QY 1128 AAAGCCTGGCTCTATCACACCCGACACCAAGTTTGGAGCTTTGTGTATGCTTAGAA 1187

Db 361 AAAGCCTGGCTTCCATCAGCGCACACCAAGTTTGGAGCTTTGTATGTGCTCAAGAA 420
 QY 1188 GGAAGAGGTGGCGGACACACTACCTTTCTTCCCTGTTACCGCCACAGTTCTACATGCG 1247
 Db 421 GGAGAGGGTGGCGGACACTCCCATTTTCTCTGTTACCGTCCGCAATTTCTACATGCG 480
 QY 1248 GACAACTGATGTGACAGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGAA 1307
 Db 481 GACTACTGATGTACCGGGAATGTGACAACATATGAATGACAAGGATGAGGAGCGAA 540
 QY 1308 GATGTGATGCTGCTGGTGACCGTATCAAAATGATTTAGCTCATCCAGCCTTCTCTTG 1367
 Db 541 GATGTGATGCTGCTGGTGACCGTATCAAGATGTTGTTGAGCTCATCCAGCCCTGCTTG 600
 QY 1368 TGAGCAGGATGATGAGTTTCTGCTATCCGTGAGGTGGTAAAGACCTGGTCCCGTGTCTAT 1427
 Db 601 TGAGCAGGATGAGTTTCTGCTATCCGTGAGGTGGTAAAGACCTGGTCCCGTGTCTAT 660
 QY 1428 CAACAAATCATTCAGTAA 1447
 Db 661 CAATAATCATTCAGTAA 680
 RESULT 5
 AV943931
 LOCUS
 DEFINITION
 AV943931 K. Sato unpublished cDNA library, strain H602 adult,
 heading stage top three leaves Hordeum vulgare subsp. spontaneum
 cDNA clone bahl917 3', mRNA sequence.
 AV943931
 EST.
 SOURCE
 ORGANISM
 Hordeum vulgare subsp. spontaneum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 736)
 Sato, K., Saisho, D. and Takeda, K.
 Barley EST sequencing project in NIG and Okayama Univ
 Unpublished (2002)
 Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1. 736
 /organism="Hordeum vulgare subsp. spontaneum"
 /strain="H602"
 /db_xref="taxon:77009"
 /clone="bahl917"
 /clone_lib="K. Sato unpublished cDNA library, strain H602
 adult, heading stage top three leaves"
 /tissue_type="top three leaves"
 /dev_stage="adult, heading stage"
 189 a 214 c 158 g 173 t 2 others
 BASE COUNT
 ORIGIN
 Query Match 31.1%; Score 523.4; DB 10; Length 736;
 Best Local Similarity 86.1%; Pred. No. 2.5e-94;
 Matches 617; Conservative 0; Mismatches 91; Indels 9; Gaps 3;
 QY 903 CTCCTATCAGCGTCTGGTGTACAGTTCCACTGCGCGTATAGACCGTGGCACCGTCAAGAT 962
 Db 736 CTCCTATCAGCGTCTGGTGTACAGTTCCACTGCGCGTATAGACCGTGGCACCGTCAAGAT 677
 QY 963 TGTGTACACAGTGTATGCTCGGAATCCGGACACCCGGAACCTGACCGTCTACTGGTGT 1022
 Db 676 TGGGGACCCAGTGTGACCTCTGCGCATCAGGAGAGAGCCCAATGCGCAGTCTACTGGTGT 617


```

DEFINITION BJ308303 Y. Ogihara unpublished cDNA library, Wh_yd Triticum
aestivum cDNA clone whyd15g14 3', mRNA sequence.
ACCESSION BJ308303
VERSION BJ308303.1 GI:20116830
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 728)
Ogihara.Y. and Murai.K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..728
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whyd15g14"
/tissue_type="spikelet at late flowering"
/dev_stage="Feekes' scale 6"
/Note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J. Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give phagescript phagemids in the TJ Close lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."
BASE COUNT 194 a 212 c 153 g 169 t
ORIGIN
Query Match 30.9%; Score 520; DB 13; Length 728;
Best Local Similarity 86.3%; Pred. No. 1.2e-93;
Matches 635; Conservative 0; Mismatches 90; Indels 11; Gaps 5;
QY 881 TGCTGCTGTTCAAGATGCTTCT-CCATCACCGGTCTGTGTACAGTTCGCACTGG-CCG 938
Db 728 TGCTGCTGTTGAGGATGCTTCTCCCATCACTGTGTGTACAGTTCGCACTGGCCG 669
QY 939 TATAGAGGTGGCCACCGTCAAGATTGGTGACACAGTCATATCTCGGAATCCGGGACAC 998
Db 668 TATCAGCGTGGCACCCTGAGTTGGGACCCAGTCCACCTCGTGGCATCAGGGAGAC 609
QY 999 CCGGAACCTGCAGGTGCTACTGTGTGAGATGTTCCAGAAGACCATGGATGCCATGGC 1058
Db 608 TCGCAATCCACCGGTCACTGTGTGTGAGATGTTCCAGAAGACCATGGATGCCATGGC 549
QY 1059 CGGAGACAATGTGGGTGCTGCTCCGTGGTATGCACGAAGATGATGATGATGATGATG 1118
Db 548 TGGGACAATGTGGGTGCTGCTCCGTGGTATGCAGGAAGAGACATTTGAGAGAGCAT 489
QY 1119 GGTGCTGCAAGAGCTGCTCTATCACACCGCACACCAAGTTTGAGGCTGTGTGATGT 1178
Db 488 GGTGTGGCAAGAGCGGTTCATATCACCGCACACACCAAGTTTGAGGCTGTGTGATGT 429
QY 1179 GCTTAAGAAGAGAGGGTGGCCACACTCACCTTCTCCCTGGTTACCGCCACAGTT 1238
Db 428 GCTCAAGAAGAGAGGGTGGCCGACCTCCCATTTTTCCTGGTTACCGTCCGCAATT 369

```

```

QY 1239 CTACATGGCGGACAACTGATGTGACAGGAGTGTGACTACGATTATGAATGACAAGATGA 1298
Db 368 CTACATGGCGGACTACTGATGTACGGGAATGTGACAAACATTATGAATGACAAGATGA 309
QY 1299 GGAGCGGAAGATGTGCATGCCCTGGTGACCGTATCAAAATGATTGTTCACTCATCCAGCC 1358
Db 308 GGAGCGGAAGATGTGCATGCCGGGTGACCGTATCAAGATGTTGTGGAGCTCATCCAGCC 249
QY 1359 TGTTCCTTTGTGAGCAGGATGATGAGGTTTGTCTATCCGTGAGGTGGTAAGACCGTTGGTGC 1418
Db 248 CGTGGCTTTGTGAGCAGGGAATGAGGTTTGCCATCCGTGAAGTGGCAGACCGTGGTGC 189
QY 1419 CGGTGTCATCAACAATAATCATTTAGTAACTGGATATAACATATCCACCATGAGAATTTT 1478
Db 188 CGGTGTCATCAATAATATCATTTAGTAACTGGATATAACATATCCACCGTGAAGAATTTT 129
QY 1479 CCTTGTCTTACTC-AAAGCGACATGCTCGTAGTTGTTTATTTATGTTGTTAGGGGT 1537
Db 128 CCTCATTTACTCTTTTGGGAATGCTATGTTGTTATTTATG-----CATTTAGGGA 75
QY 1538 TGCTCATGTGCAATTTAGTATGACACTTTTTTTTGTCAAGTGAATTTGCATATTTAT 1597
Db 74 GGCTCTTTGAAATTTAGTATGACACATTTTCT--TCAGTGAATTTGCATCTTTGT 17
QY 1598 GACATTCACGACAAAG 1613
Db 16 AGTATTCACGACAAAG 1
RESULT 8
BG301275/c
LOCUS BG301275 900 bp mRNA linear EST 22-OCT-2001
DEFINITION HVSMEB0020G16f Hordeum vulgare seedling shoot EST library
HVCDNA0002 (Dehydration stress) Hordeum vulgare cDNA clone
HVSMEB0020G16f, mRNA sequence.
ACCESSION BG301275
VERSION BG301275.1 GI:13098802
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 900)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
,R.D., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex drought-stressed seedling shoot cDNA
library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 221
Seq primer: AATTAACTCACTAAAGGG
High quality sequence start: 8
High quality sequence stop: 739.
FEATURES
Location/Qualifiers
1..900
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEB0020G16f"
/HVCNA0002 (Dehydration stress)
/tissue_type="Seedling shoot"
/lab_host="TJJC121"

```

/note="Vector: lambdaZAP; Site.1: EcoRI; Site.2: XhoI; Seeds were surface sterilized then germinated under axenic conditions in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Five-day old seedlings were incubated at 90% RH for 24 hr. Shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, 600000 pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT
ORIGIN

225 a 258 c 212 g 205 t
Query Match 30.7%; Score 515.4; DB 12; Length 900;
Best Local Similarity 80.1%; Pred. No. 9.4e-93;
Matches 707; Conservative 0; Mismatches 161; Indels 15; Gaps 8;

QY 734 GCTCCGCTCAGGCGCTGAGGCTCTCATGCTCAACCTCGCTTGAAGCGGGGAGC 793
DB 874 GGATGACCTCAAGGCGCTGAGTCTTCAT-GCCACCCTGGCTCCAAGGGTGGGAATA 816
QY 794 ATGAGTGGGCGACTACATCTCTCGTTGGTGTGATAAAGTGGATTCCTATATCCAGTCC 853
DB 815 AAGAGTGGGTTGAGGCG-TCCTTTCCCTGATGCTCCCTGACCACCCAAAT--CCTGTC 759
QY 854 CGCAGAGGCGAGCTGACCTCCCTCTTCTGCTCGCTGTTGAAGATGCTTCTCCATCACCG 913
DB 758 CGCAAGGCGAGACAGACCTGCCCTTC-TGTTCTGCTGTAGGATGCTTTTCTCTCCCG 700
QY 914 GTCGTGGTAC-AGTTGCCACTGCCCTATAGACCTGGCCACCGTCAAGATTTGTCACACA 972
DB 699 GTCGGGTCCCAAGTTCACCTGACCTGACCTGACGCGTGGCAGCGTCAAGTTGGGGACCA 640
QY 973 GTCGATATCTCGGAATCCGGGACACCGGAACTGCACCGTCACTGGTGTGAGATGTTTC 1032
DB 639 GTCGACCTCTCGGCATCAGGAGACCGCAATGCCAGTCACTGCTGTAGATGTTTC 580
QY 1033 CAGAAGACCATGATGATGCCATGCGCGGAGACAAATGTTGGCTGCTGCTCGCTGGTATG 1092
DB 579 CAGAAGACCATGATGATGATGCTGTTGGGACAAATGTTGGCTGCTGCTGCTCGCTGATG 520
QY 1093 CAGAAGATCACATGAAAGGCGATGCTGCTGGCAAGCCCTGGCTCTATCACACCGCAC 1152
DB 519 CAGAAGAACACATGAGATAGGATGCTGTTCTCAAGCCGGGTTCATCATCGCCACAC 460
QY 1153 ACCAAGTTTGGAGCTGTTGTATGCTTAAGAAGAAAGAGGGTGGCGGCACACTCACCT 1212
DB 459 ACCAAGTTTGGAGCTGTTGTATGCTCAAGAAGAGGAGGGTGGCGGCACCTCCCA 400
QY 1213 TTTCTCCCTGGTTACCGCCACAGTTCTACATCGCGACAACTGATGTGACAGGAGTGTG 1272
DB 399 TTTTCCCTGGTTACCGCTCGGACAGTTCTACATCGCGACTACTGATGTCAGGGGACGCTG 340
QY 1273 ACTACGATTTATGACAGGATGAGGAGCGGCAAGATGTGATGCTGCTGACCGCTATC 1332
DB 339 ACAACATTTATGACAGGATGAGGAGCGCAATATGTGATGCTCGCGGTGACCGCTATC 280

QY 1333 AAAATGATTGTTGAGCTCATCAGCCTGCTCTGTTGTGAGCAGGTGATGAGTTCCTGCTATC 1392
DB 279 AAGATGGTTGGAGCTCATCAGCCTGCTGCTGTTGTGAGCAGGAAATAGTTCCTGCTATC 220
QY 1393 CGTGAAGTGGTGAAGACGCTGTTGCGCGTGTCTATCAACAAATCATTTAGTAACTGGA 1452
DB 219 CGTGAAGACGCGACAGACAGTCCGTGCTCATCAACAAACATCATTCAGTAAAGTGA 160
QY 1453 TATAACATATCCACCATGAGATTTTCCTGTTTACTC-AAAGCGACATGCTCGTAGTT 1511
DB 159 TGGAGATATCCCGCTGAGAAATTTTCCTCATTTGCTGACTCTTTTGGGAAATGCTATGAT 100
QY 1512 GTTATTATGTTGGTGGTGTGCTCATGTGCAATTTGTAGTATGACACATTTTTTTT 1571
DB 99 GCTATTATG-----GGTTTAGGAGAGCTCTTGTGAATTTGTAGTATGCTCTTCTTC 46
QY 1572 TTGCAACTGAATTTTGCATAATTTTATGACATTCACGACAAAGA 1614
DB 45 T--TCAAGTGACATTGCAAACTTTGTAGTATTCGCGGCAAAA 5

RESULT 9
BJ263237/c

LOCUS BJ263237 Y. Ogiwara unpublished cDNA library, wh_h Triticum
DEFINITION aestivum cDNA clone whh13j11 3', mRNA sequence.

ACCESSION BJ263237
VERSION BJ263237.1 GI:20084049
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 685)
AUTHORS Ogiwara, Y. and Murai, K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..685
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whh13j11"
/clone_lib="Y. Ogiwara unpublished cDNA library, wh_h"
/tissue_type="spike at heading date"
/dev_stage="Feekes' scale 10.5"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site.1: EcoRI; Site.2: XhoI; plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J. Dvorak Lab). total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pBluescript phagemids in the TJ Close lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."

BASE COUNT 179 a 201 c 144 g 159 t
ORIGIN

Query Match 30.3%; Score 509.6; DB 13; Length 685;
Best Local Similarity 86.4%; Pred. No. 1.4e-91;
Matches 598; Conservative 0; Mismatches 86; Indels 8; Gaps 3;

RESULT 11

BI959687

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

BI959687 678 bp mRNA linear EST 22-OCT-2001

HVSME0020M20f Hordeum vulgare rachis EST library HVCDNA0015

(normal) Hordeum vulgare cDNA clone HVSME0020M20f, mRNA sequence.

BI959687

BI959687.1 GI:16310942

EST

Hordeum vulgare.

Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

; Triticeae; Hordeum.

1 (bases 1 to 678)

Wing, R., Close, J., Klein, R., Wise, R., Chin, A., Begum, D.,

Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons

J., Oates, R. and Main, D.

Development of a genetically and physically anchored EST resource

for barley genomics: Morex rachis cDNA library

Unpublished (2001)

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: twing@clemson.edu

Total hg bases = 490

Seq primer: AATTACCTCCTCACTAAAGG

High quality sequence stop: 556.

Location/Qualifiers

1..678

/organism="Hordeum vulgare"

/cultivar="Morex"

/db_xref="taxon:4513"

/clone="HVSME0020M20f"

/clone_lib="Hordeum vulgare rachis EST library HVCDNA0015

(normal)"

/tissue_type="rachis"

/lab_host="TJCI21"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; Plants were grown at Washington State University,

Pullman, WA in a greenhouse, the rachises were excised and

frozen in liquid nitrogen (Klein, R. and Main, D.). In the TJ Close

lab at the University of California, Riverside total RNA

was prepared, poly(A) was purified, one primary

unamplified cDNA library was made, and 1 million pfu were

in vivo excised to give pBluescript SK(-) cDNA phagemids

(Chin). Phagemids were plated and picked at the Clemson

University Genomics Institute (CUGI) (Begum, Palmer, M.

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA

sequencing and sequence analysis were performed at CUGI

(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The

sequence has been trimmed to remove vector sequence and

contains a minimum of 100 bases of phred value 20 or

above. For more details on library preparation and

sequence analysis see

http://www.genome.clemson.edu/projects/barley. To order

this clone see http://www.genome.clemson.edu/orders Also

see Close TJ, Wing R, Klein, R., Klein, R. and Wise R (2001)

Genetically and physically anchored EST resources for

barley genomics. Barley Genetics Newsletter 31:29-30.

(http://wheat.fw.usda.gov/gpages/bgn/31/cover.html)"

BASE COUNT 131 a 236 c 188 g 123 t

ORIGIN

Query Match 29.8%; Score 501.4; DB 13; Length 678;

Best Local Similarity 86.1%; Pred. No. 6.2e-90;

Matches 578; Conservativity 0; Mismatches 91; Indels 2; Gaps 2;

QY 247 CCACGTCACATAGGACCATCGGCATGTGCGACACGAGAAAGACCCTCTCACC

Db 1 CCACAGTCACACCGCGGACCATCGGCACATCGGCACGCGGACGCGGACGCGGACGCGG 60
 QY 307 GGGCTCACATAGTGTCTCGCTCCCTCGGTGCGAGCGGCGCTAGAAAGTACGACGAGATC 366
 Db 61 GGGCTCACATAGTGTCTCGCTCCCTCGGTGCGAGCGGCGCTAGAAAGTACGACGAGATC 120
 QY 367 GAGCGCGCGCGGAGGAGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAGTAC 426
 Db 121 GAGCGCGCGCGGAGGAGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAGTAC 180
 QY 427 GAGCGCGCGCGGAGGAGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAGTAC 486
 Db 181 GAGCGGAGAGCGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAGTAC 240
 QY 487 AATATGATCATCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 546
 Db 241 AATATGATCATCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 300
 QY 547 GGGCGCATGCGCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 606
 Db 301 GGGCGCATGCGCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 360
 QY 607 ATCGTGTGCTTCTCTCAACAAGAGGACATGCTCGGCGGCGGCGGCGGAGTAC 666
 Db 361 ATCGTGTGCTTCTCTCAACAAGAGGACATGCTCGGCGGCGGCGGCGGAGTAC 420
 QY 667 GAGCTCGAGGTCGCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 726
 Db 421 GATCTCGAGGTCGCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 480
 QY 727 GTCGCTGGCTCGCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 786
 Db 481 GTCGCTGGCTCGCGGCGGCGGCGGCGGCGGTATCACCATCAACACGCGGCGGCGGAG 540
 QY 787 GCGCGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 846
 Db 541 GGGGATAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 QY 847 CCAGTCCCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 904
 Db 601 CCGTACCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 660
 QY 905 CCATCACCGGT 915
 Db 661 CCATCACCGGT 671

RESULT 12

AV938814

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 644)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

AV938814 644 bp mRNA linear EST 18-JAN-2002
 AV938814 K. Sato unpublished cDNA library, strain H602 adult,
 heading stage top three leaves Hordeum vulgare subsp. spontaneum
 cDNA clone bahl917 5', mRNA sequence.

AV938814
 EST
 Hordeum vulgare subsp. spontaneum.
 Hordeum vulgare subsp. spontaneum

AV938814
 EST
 Hordeum vulgare subsp. spontaneum.
 Hordeum vulgare subsp. spontaneum

AV938814
 EST
 Hordeum vulgare subsp. spontaneum.
 Hordeum vulgare subsp. spontaneum

AV938814
 EST
 Hordeum vulgare subsp. spontaneum.
 Hordeum vulgare subsp. spontaneum

AV938814
 EST
 Hordeum vulgare subsp. spontaneum.
 Hordeum vulgare subsp. spontaneum

Db	69	CTCGTGGGGCAGCGGCCCAAGAAGTACGACGAGATCGAGCGCCCTCGAGGAGCG	128
QY	387	CGCCGGGGTATCACCATCAACACCGCCACCGTCTGAGTACGAGACGAGACCGCCCACTA	446
Db	129	CGCCGGTGGTATCACCATCAACACCGCCACCGTCTGAGTACGAGACGAGACCGCCCACTA	188
QY	447	CGACACGTGACATGCCCGCCGACGCGGACTATGTCAAGAATATGATCACGGCGCTCG	506
Db	189	CGCGACGTGACATGCCCGCCGTCACGCGGACTACGTCAAGAATATGATCACGGCGCTCG	248
QY	507	GCAGATGCAGCGGTGCCATCTCTGCTATCCGGTGCAGCGGCGCCATGCCGACACCAA	566
Db	249	CCAGATGCAGCGGCGCCATCTCTGCTCTCCGCGCGGAGCGGCCCATGCCGACACCAA	308
QY	567	AGACACATCTCTCGCCCAAGCAAGTCGTGTTCACAGATGCTTCTCTCTCAACAA	626
Db	309	GGAGCACATCTCTCGCCCAAGCAGGTGCGTGTCCCGCAGCATTTGTTCTCTCAACAA	368
QY	627	GAAGGACATGGTCACGACGAGGAGTGTCTCGAGTCTGTCGAGTCTCGAGGTCGGGAGCT	686
Db	369	GAAGGACAGTGCAGCAGCAGAGCTGCTCGAGTCTGTCGAGTCTCGAGGTCGGGAGCT	428
QY	687	GCTCAGCAATCAGAGTACGACGGCGACGACGTACCAATCGTGTGCTCGGCTCGCCCTCAA	746
Db	429	GCTCAGCGCTATCAATACGATGCTGACAAATGTGCCAATCGTCTCGGCTCTGCACTCAG	488
QY	747	GGGCTCGAGGCTCTCATGTCAACCGTGCCTTGAGCGCGGCGGACGATGAGTGGGTGA	806
Db	489	AGCGCTCGAGGCGCTCATGCGCCACCGCTGCGCTCAAGCGTGGGGATAACGAGTGGGTGA	548
QY	807	CTACATCTTCTCGTGTGGTATAAAGTGGATTCCTATATTCAGTCCCGCAGGACGAG	866
Db	549	TGGCATCTTCTCTGATTGATCCGTGGATACCCACATCCCTGTCCCGCAGGACGAG	608
QY	867	TGACCTCCCGTCTTCTGCTGCTGT	890
Db	609	CGACCTCCCGTCTTCTGCTGCTGT	632
RESULT 15	BI992656/C		
LOCUS	1020065H09.x1 1020 - Unigene II from Maize Genome Project	558 bp mRNA linear	EST 24-OCT-2000
DEFINITION	cdna, mRNA sequence.		
ACCESSION	BI992656		
VERSION	BI992656.1	GI:16378882	
KEYWORDS	EST.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC		
AUTHORS	clade; Panicoidae; Andropogoneae; Zea.		
TITLE	1 (bases 1 to 558)		
JOURNAL	Walbot.V.		
COMMENT	Maize ESTs from various cdna libraries sequenced at Stanford University Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 1020065 row: H column: 09.		

for the maize genome project. Sequences are present from libraries 947 and 949. Contigs were assembled using zmbassembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

BASE COUNT	156 a	154 c	106 g	142 t
ORIGIN				
Query Match	29.3%; Score 493.2; DB 13; Length 558;			
Best Local Similarity	93.8%; Pred. No. 2.8e-88;			
Matches	525;	Conservative	0;	Mismatches 33; Indels 2; Gaps 1;
Qy 1048	GATGCCATGGCGGAGACAATGTTGGCTGCTGCCGTGATGACAGAGGATGACATT 1107			
Db				
Qy 558	GATGCCATGGCTGGAGACAATGCTGGACTGCTGCCGTGATGACAGAGGATGACATT 499			
Db				
Qy 1108	GAAAGAGGCATGGTGCCTGGCAAGCCCTGGCTCTATCACCGCACACCAAGTTGAGGCT 1167			
Db				
Qy 498	GAGAGAGGCATGGTGCCTGGCAAGCCCGGCTCTATCACCGCACACCAAGTTGAGGCT 439			
Db				
Qy 1168	GTTGTGATGTCCTTAAGAAGAGAGGGTGGCGGACACTCACCTTCTTCCCTGGTTAC 1227			
Db				
Qy 438	GTTGTGATGTCCTTAAGAAGAGAGGGTGGCGGACACTCACCTTCTTCCCTGGTTAC 379			
Db				
Qy 1228	CGCCACACAGTTCTACATCGGACAACTGATGTGACAGGGAGTGTGACTACCAATTATGAAT 1287			
Db				
Qy 378	CGCCGCGCAGTTCTACATGAGCACAAACCGATGTGACAGGAATGTGACTGTGATTATCAAT 319			
Db				
Qy 1288	GACAAAGATGAGGAGGCGAAGATGTGCATGCTGCTGACCGTATCAAAATGATTGTTTAC 1347			
Db				
Qy 318	GACAAAGATGAGGAGGCGAAGATGTGCATGCTGCTGACCGTATCAAGATGGTTGTTTAC 259			
Db				
Qy 1348	CTCATCCAGCCTGTTGCTTGAGCAGGGTATGAGGTTTGTCTATCCGTGAGGGTGGTAAG 1407			
Db				
Qy 258	CTCATCCAGCCTGTTGCTTGAGCAGGGTATGAGGTTTGTCTATCCGTGAGGGTGGTAAG 199			
Db				
Qy 1408	ACCGTTGGTGGCGGTGTCATCAACAAATCATTTGAGTAACTGGATATACATATCCACC 1467			
Db				
Qy 198	ACTGTTGGTGGGAATAAACAACAAATCATTTGAGTAACTGGATATACATATCCACC 139			
Db				
Qy 1468	ATGAGAAATTTCCCTGTTTACTCAAGCGACATGCTCGTAGTTGTTATTATGTTGAG 1527			
Db				
Qy 138	ACGAGAAATTTCCCTGTTTACTCAAGCGAATGCTCTGTAGTTGTTATTATGTTGAG 79			
Db				
Qy 1528	TTTTAGGGGTTGCTCATGTGCAATTGTAGTATGACACTTTTTTTTTTGTCAAAGTGAATTG 1587			
Db				
Qy 78	TTTTAGGGGTTGCTCATGTGCAATTGTAGTATGACAC--TCCTTTCGCCCAAGTGAATTG 21			
Db				
Qy 1588	CATAATTTATGACATTCACG 1607			
Db				
Qy 20	CATAATTTATGACATTCACG 1			

Search completed: May 1, 2003, 10:18:51
Job time : 2907 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:08:30 ; Search time 98 Seconds
(without alignments)
5260.445 Million cell updates/sec

Title: US-09-810-764A-6

Perfect score: 1681

Sequence: 1 attcccataatcccccacc.....gttaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	488.4	29.1	4411529	4	US-09-103-840A-1
2	477.4	28.4	4403765	4	US-09-103-840A-2
3	420.4	25.0	1185	2	US-08-743-637B-185
4	402.8	24.0	1245	4	US-09-134-001C-2540
5	393.2	23.4	1185	4	US-09-218-197-1
6	368.8	21.9	2996	4	US-08-961-527-260
7	243.8	14.5	2592	4	US-09-221-017B-642
8	188.8	11.2	1230	4	US-09-140-466-1
9	182.2	10.8	3719	1	US-08-920-812-10
10	182.2	10.8	3719	1	US-08-920-827-10
11	182.2	10.8	3719	1	US-08-921-177-10
12	182.2	10.8	3719	1	US-08-362-570-10
13	182.2	10.8	3719	2	US-08-920-828-10
14	180.2	10.7	656	4	US-08-936-165A-87
15	112.6	6.7	323	4	US-09-556-877-70
16	112.6	6.7	323	3	US-09-620-412C-70
17	105.8	6.3	785	3	US-09-154-083-29
18	77.8	4.6	390	4	US-09-197-649-7
19	71.4	4.2	30001	1	US-08-125-468-1
20	71.4	4.2	30001	2	US-08-474-933-1
21	70.2	4.2	2095	1	US-08-329-681A-2
22	63	3.7	1248	4	US-09-105-537-7
23	63	3.7	13613	4	US-09-105-537-3
24	62.6	3.7	546	4	US-09-194-905-3
25	62.2	3.7	5970	3	US-09-320-878-21
26	62.2	3.7	6854	4	US-09-194-905-7
27	61.6	3.7	3461	2	US-08-389-564B-1

28	61.6	3.7	3461	3	US-08-466-047B-1	Sequence 1, Appli
29	61.4	3.7	1500	4	US-09-174-768-3	Sequence 3, Appli
30	60.6	3.6	390	4	US-09-197-649-7	Sequence 7, Appli
31	58.6	3.5	12588	2	US-08-387-942C-1	Sequence 1, Appli
32	58.4	3.5	8438	1	US-07-945-283-1	Sequence 1, Appli
33	58	3.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli
34	58	3.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli
35	57.6	3.4	1896	4	US-09-311-626B-15	Sequence 15, Appli
36	57.4	3.4	1521	1	US-08-083-948-14	Sequence 14, Appli
37	57.4	3.4	1521	1	US-08-393-785-14	Sequence 14, Appli
38	57.4	3.4	1521	1	US-08-475-694-14	Sequence 14, Appli
39	57.4	3.4	1521	1	US-08-712-057-14	Sequence 14, Appli
40	57.4	3.4	1647	1	US-08-083-948-13	Sequence 13, Appli
41	57.4	3.4	1647	1	US-08-393-785-13	Sequence 13, Appli
42	57.4	3.4	1647	1	US-08-475-694-13	Sequence 13, Appli
43	57.4	3.4	1647	1	US-08-712-057-13	Sequence 13, Appli
44	57.4	3.4	1865	1	US-08-083-948-7	Sequence 7, Appli
45	57.4	3.4	1865	1	US-08-393-785-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match	29.1%	Score	488.4;	DB	4;	Length	4411529;
Best Local Similarity	65.8%	Pred. No.	1.5e-103;				
Matches	815;	Conservative	0;	Mismatches	366;	Indels	57;
Gaps	57;						
QY	217	GC	GGCGAGGGGCAAGTTCGAGCGGCACCAACACACAGTCAACATAGGACCATCGGCCAT	276			
Db	784819	GT	GGCGAAGGCGAAGTTCACGCGGACCAAGCCCAACATCGGACCATCGGTAC	784878			
QY	277	GT	CGACACCGAAGAACCACTCTCAGCGCGCTCACCATGGTGTCT	330			
Db	784879	GT	TGACACCGGCAAGACCACTCTGACCGGCTATCACCAGGTCTGACGACCAATTC	784938			
QY	331	GT	CGGTGGAGCGCGCTTAAGAGTACGACGAGATCGACCGCGCCCGGAGGAGCGCGCC	390			
Db	784939	CC	GGTCTGAACGAGAGCAAGGATTCGACCAAGCCCGGAGGAGCGGTAC	784998			
QY	391	CG	CGGTATCACCATCAACACCGCCCGCTGAGTACGACCGGACCGCCCATACGCA	450			
Db	784999	CG	CGGTATCACCATCAACATCGCGCAGTGGAGTACGACCGGACCAAGCGGTACGCA	785058			
QY	451	CAG	CTCGACTGCCCGCGCCCGGCTATGTCAAGAAATATGATCACCAGGTCTGCGCAG	510			
Db	785059	CAG	CTCGAGCGCCCTGGCCACCGCGCTATCATCAAGACATGATCACCAGCGCGCGCAG	785118			
QY	511	AT	GACGGTGCATCTCTGCTGATCCGGTGGCGGCGCCATCGCGGACCAAGAG	570			
Db	785119	AT	GACGGTGCATCTCTGCTGATCCGGTGGCGGCGCCATCGCGGACCAAGAG	785178			

Db	787527	TCGCCATCGACCACACGACCGTCACCGGTGGAGATGTTCCGCAAGCTGCTCGA	787586
Qy	1047	TGATGCCATGCCCGGAGACAATGTTGGCTGTGCTCGGTGATGCGAAGGATGACAT	1106
Db	787587	CCAGGCGACGCGGGGACACAGTTGGTTGCTGCGGGCGTCAAGCGCGAGGACGT	787646
Qy	1107	TGAAGAGGCGATGGCTGTGCAAGAGCGCTGGCTCTATCACACCGCACACAAAGTTTGAGGC	1166
Db	787647	CGAGCGTGGCCAGTGTGCACCAAGCCCGCACACCGAGTTCGAAGG	787706
Qy	1167	TGTTGTGTATGTGCTTTAAGAAGAGAGGTTGGCGACACTCACCTTTCTCCCTGGTTA	1226
Db	787707	CCAGGCTACATCCTGTCCAAAGGACGAGGCGCGGACACGCGTTCCTCAACAAC	787766
Qy	1227	CCGCCACAGTTCTACATCGGCACAACTGATGTAGCAGGAGTGTGACTAGCATTTATGAA	1286
Db	787767	CCGTCCGCGAGTTCTACTTCCGACACCACCGACGTGACCGTGTGGTGAC-----	787814
Qy	1287	TGACAAGGATGAGGAGCGGAAGATGTCATGCTCGTGACCGTATCAAATGATTGTTC	1346
Db	787815	---ACTGCGGAGGCGACCGAGATGGTATGCCGTGACACACCAACATCTCGGTGAA	787871
Qy	1347	GCTCATCCAGCCTGTTGCTGTGACGAGGATGAGGTTTGTATCCGTGAGGTTGTTAA	1406
Db	787872	GTTGATCCAGCCGTCGCCATGGAGAGGTCGTGCTTCGGATTCGCGAGTCCGCGAGGTGGCG	787931
Qy	1407	GACCGTTGGTCCGGTGTCTATCAACAAATCATTTAGTA	1445
Db	787932	CACCGTGGCGCGCGCGGTTCACCAAGATCATCAAGTA	787970

RESULT 3

US-08-743-637B-185
; Sequence 185, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 185:
; SEQUENCE CHARACTERISTICS:

Db 80 CCTAAAGACTATGGCTCTATCGATGCTCTCCAGAGAAACCGGACGCTATCACTATC 139
Qy 406 AACACGGCCACCGTCGAGTACGAGACGAGACCGCCACTACGACACGCTGCACTGCC 465
Db 140 AACACTGGCAGCTTGAGTAGCAAACTGAAAAGCTCACTACGCTACATCGAGCTCCA 199
Qy 466 GGCACCGCCGACTATGTCAAGAATATGATCACCGCGCTCGCGAGATGGAGCGTGCATC 525
Db 200 GGACACGGGACTACGTTAAACACATGATCACTGGTCTGCTCAATGGACGGAGCTATC 259
Qy 526 CTCGTCGATCCGTCGGGAGCGGCCCATGCCGCGAGACCAAGAGACACATCTCTCCGCC 585
Db 260 CTTTGTAGTGTTCACCTGAGCGACCAATGCCAATACTCGTGAGCACATCTCTCTTCA 319
Qy 586 AAGCAAGTCGCTGTCCCAAGATGTTGCTCTCTCAACAAGAGGACATGGTCGAGCAC 645
Db 320 CGTCAGGTGGTGTAAACACCTTATCGTCTTCAAGAACAAAGTTGACTTGGTTGAGCAC 379
Qy 646 GAGGAGCTGCTCGAGCTCGTGCAGTCCGAGTCCGCGAGCTGCTCAGCAACTACGAGTAC 705
Db 380 GAAGAATTGCTTGAATTGTTGAATGGAATCGGTGACCTATTGTCAGAATACGACTTC 439
Qy 706 GACGGCAGCAGTACCAATCGTCGCTCGCTCCGCTCAAGCGCTCGAGGCTCTCATG 765
Db 440 CCAGTGACGATCTTCCAGTTATCCAAAGTTTACGACTTAAAGCTCTTGAAGGT----- 493
Qy 766 GTCRAACCTGCTTGAAGCGGGCGAGATGAGTGGTGCAGTACATCTTCTCGTTGGTT 825
Db 494 -----GACTCTAAATACGAAGACATCGTTATGGAATTGATG 529
Qy 826 GATAAAGTGGATTCTATATTCAGTCCCGCAGAGCGAGACTGACCTCCCGTTCTTGCTC 885
Db 530 AACACAGTTGATGATATATCCAGAACCCAGAACGTGACACTGACAAACCATGCTTCT 589
Qy 886 GCTGTTGAAGATGCTTCTCCATCACCGCTGGTGTACAGTGGCCACTGCGCGGTATAG 945
Db 590 CCAGTCGAGGACGATTCTCAATCACTCGAGCTGGTACAGTCTTCCAGGACGATCGAC 649
Qy 946 CTGCGACCGTCAAGATGGTGCACAGTGCATATCGTCGGAATC---CGGACACCCGG 1002
Db 650 CTGCTATCGTTAAGTCAAGCAGCAATCGAATCGTTGATCAAGAGAACTCAA 709
Qy 1003 AACTGCACGGTCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1062
Db 710 AAAGCAGTGTGTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 769
Qy 1063 GACATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122
Db 770 GATACGAGTGTGCTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 829
Qy 1123 CTGCAAGCGCTGCTCTATCACACCGCAGACACCAAGTTTGGAGTGTGTTGTTGTTGTT 1182
Db 830 ATCGTAAACACGTTCAATCAACCCACACACTAAATCAAGGTGAAGTCTACATCCTT 889
Qy 1183 AAGAGGAAGAGGTGGCGGACACTACCTTCTTCCCTGTTTACCAGCCACAGTTCTAC 1242
Db 890 ACTAAGAAGAAGGTGGAGCTGACACTCCATCTTCAACAACACTACCGTCCCAATTTCT 949
Qy 1243 ATGGGCAACTGATGTGACAGGAGTGTGACTACGATTATGATGACAGGATGAGGAG 1302
Db 950 TTCCGTACTGACTGTTACAGG-----TTCAATCGAATCTCCAGCAGGT 994
Qy 1303 GCGAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
Db 995 ACTGAATGGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
Qy 1363 GCTGTGAGCAGGTATGAGTGTGCTATCGCTGAGGCTGTTAGACCGTGTGGCGGTT 1422
Db 1055 GCGGTAGAACAAGGTACTACATTTCTTCTATCGTGGGTTGAGCTGCTGTTGTTGAGT 1114
Qy 1423 GTCATCAACAAATCAATTTGAGTAACTGGATATAACATATCCACATGAGAAATTTTCC 1482
Db 1115 ATGGTTACAGAAATCGAAGCTTAATTCGATTTAGTTCCTCCAGAACAAATATTTAAGT 1174

Qy 1483 GTTACTAAAGCAGCATGCTCCGTAGTTGTTATTTATTTATGTTGAGT 1528
Db 1175 AGACACTAAAGAATCTTGCTGGCAAGGTTCTTTTATTAGATATT 1220

RESULT 7

US-09-221-017B-642
; Sequence 642, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 642:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2592 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...2592
US-09-221-017B-642

Query Match 14.5%; Score 243.8; DB 4; Length 2592;
Best Local Similarity 60.9%; Pred. No. 1e-47;
Matches 432; Conservative 0; Mismatches 247; Indels 30; Gaps 1;
Qy 219 GCGAGGGGCAAGTTCTGAGCGCACCAACACCACTAGGACCATCGGCATGT 278
Db 1914 GCGAAGAGCATTTTAACAGATCGAACCACCCACGTTACGTTGGTACGACGCT 1973

Qy 279 CGACCACGAAAGACCACTCTCACCAGCGGCTCACCATGGTGTGCTCCCTCCGTCGGTGG 338
Db 1974 GGACCAGGTAAGACTACCTTGACGGCTGCAATCAACAACCTGTGTGCAAGAGGGGTCT 2033
Qy 339 CAGCGCGCTAAGAGTACGACGAGATCGACGGCGCCCGGAGAGCGGCCCGCGGTAT 398
Db 2034 TTCAGAACTCCGTTTCATTTGATTCAATGATGAAGCTCCCGGAAGAAAGAAAGCTGTGTAT 2093
Qy 399 CACCATCAACACCCCGCTGAGTACGAGACCGGAGACCGGCTACGACACACGTCGA 458
Db 2094 CACGATTAACTTCACAGCTTGAATACCAACAGCTAATCGCTACGCTCACGCTGAG 2153
Qy 459 CTGCCCCGGCCACCGCACTATGTCAGAAATATGATCACCGGGCTCGCGAGATGGAAGG 518
Db 2154 CTGTCCGGGTCAACCGCTATGTCAGAAACATGTTACCGGTGCTGCTCAGATGAGCG 2213
Qy 519 TGCCATCTCGTGTATCCGGTCGCGAGCGGCCCATGTCGCGACACCAAGAGACATCTT 578
Db 2214 TGCTAATCGTTGTAGCAGCTACAGAGCGGCCCTATGCCCTCAGACTCGCGAGCACATCTT 2273
Qy 579 CTGCGCAAGCACTCGGTGTTCCCAAGATCGTTGCTCTCAACAAGAGGACATGCT 638
Db 2274 TTTGGCTCGTCAGTAAAGCTTCTCTGCTGGTGTGTTTCATGACAAATGTCACATGGT 2333
Qy 639 CGACGACGAGAGCTGCTCGAGCTGCTGAGCTCGAGGTCGCGAGCTGCTCAGCAACTA 698
Db 2334 AGACGATGAGAGATGCTCGAGCTGTTGAAATGGACATCGCGCAACTCTTCTTCTCTA 2393
Qy 699 CGACTACGCGGCGACGACCTACCAATCGCTCGCTGGCTCGCGCTCAAGCGCTCGAGGC 758
Db 2394 CGATTTCGATGGTGAACAATACCCCTATCATCCGCTGCTGCTGCGGCGCTTT----- 2447
Qy 759 TCTCATGCTCAACCCCTGCCTTGAAGCGCGGCGAGCTGAGTGGGTGCACTACATCTCTC 818
Db 2448 -----GAATGGAGACCCTCAGTGGGAGACAGGTGATGA 2483
Qy 819 GTTGGTTGATMAAGTGAATTCCTATATTCAGTCCCGCAGAGGACGACTGACCTCCCGTT 878
Db 2484 GCTTATGGAAGCTGTTGACAACTGGGTTCGCCCTGCCAGAGCGGATATCGACAAACCGTT 2543
Qy 879 CTTCCTCGCTGTTCAAGATGCTTCTCCATCACGCGGTGCTGACGTT 927
Db 2544 CTTCATCCCGTTGAAGACGTGTTCTCTATCACGGGTGCTGGTACGTT 2592

RESULT 8

US-09-140-466-1
; Sequence 1, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; FILE REFERENCE: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; CURRENT APPLICATION NUMBER: US/09/140,466
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: US 60/056,246
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1230)
US-09-140-466-1

Query Match 11.2%; Score 188.8; DB 4; Length 1230;
Best Local Similarity -48.5%; Pred No. 5,3e-35;
Matches 585; Conservative 0; Mismatches 612; Indels 9; Gaps 2;
Qy 239 GCACCAAAACACACAGCTCAACATAGGCACCATAGGCACCATGCGGCATGTGCACCAACGGAAGACACTTC 298
Db 23 GAAATAAACAACATATAAATTTAGGTACTATAGGGCATGTAGTATCATGTGAAAAAATACAT 82
Qy 299 TCACCCCGCGGCTCACCATGGTGTCTCGCTCGGTGGCAGCGGCTTAGAAGTACG 358
Db 83 TAACAACAGCTATATCTTTATTTAAATTTACAAGGATTATCAAAAAATATAATTTAT 142
Qy 359 ACAGATCGACCGCGCCCGCGAGGAGCGCGCGGTATCACCATCAACACCGCCACCG 418
Db 143 CAGATATTGATTCAGCTCCAGAGAAATAAAGAGTATTACAAATAAATACACACATA 202
Qy 419 TCGAGTACGAGACCGGACCGGCTACGACACATGCTGACCTGCCCCGGCCAGCGGACT 478
Db 203 TTGAATATGAACTTTAAACAACATTTGTCTCATATAGATTGTCAGGACATTCGGATT 262
Qy 479 ATGTCAAGAAATATGATCACCGGCTGCGCAGATGGCAGGTGCCATCTCGTCGTATCG 538
Db 263 ATATTAAAAATATGATTATAGGAGCCACACAATGGATATACAAATTTTAGTAAATATCTA 322
Qy 539 GTGCGACGCGGCCATGCGCCAGACCAAAAGACACATCTCTCGCCAAAGCAAGTCGGTG 598
Db 323 TAATAGATGGTATTAATGCTCAAACTTATGAACATTTATTATTAATAAACAATAAGSTA 382
Qy 599 TTCCCAGAGTGTGCTCTCTCAACAAGAAGACATGGTGCAGCAGGAGAGTCTGCTCG 658
Db 383 TAAAAAATAAATATTTTAAATAAAGAAGATTATGTGATGATGTTGAATTAATAG 442
Qy 659 AGTCTGCTCAGCTCGAGGTCCGCGAGCTGCTCAGCAACTACGAGTACGACGCGGACGAG 718
Db 443 ATTTATAAATTAGAAGTAATGAATTAATTAATTAATTAATTTGATTTAAATTTATA 502
Qy 719 TACCAATCGTCTGCTGCTCGCCCTCAAGCGCTCGAGGCTCTCATGTGTCACCCCTGCT 778
Db 503 TACATATATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
Qy 779 TGAAGCGCGGACGATGAGTGGTGGTACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 838
Db 563 ATGAATTAATAAATCTAATATTTGGATACAAAAATTAATAATTAATTAATTAATTAAT 622
Qy 839 CCTATTTCCAGTCCCGCAGAGCGAGCTGACCTCCGCTTCTGCTGCTGCTGCTGCTGCTG 898
Db 623 ATAAATATTATATACCTACTAGAAAAATTAATGATTACTTTTTTAATGTCATAGAGATG 682
Qy 899 TCTTCTCCATCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
Db 683 TATTTCTATAACAGGTAGAGGTACAGTAGTAACAGGTAAAGATTGAACAAGGATGTATAA 742
Qy 959 AGATGGTGACACAGTCGATATCGTCGGAAATCCGGGACACCC-----GGAAGTCCGCG 1012
Db 743 ATTTAAATGATAAATGAAATTTTAAATTTGAAAAATTCATCTCTCTAATTTTAAACAAC 802
Qy 1013 TCAGTGGTGTGAGATGTTCCAGAGAGACCATGGATGATGCCATGCCGATGAGCAATGTTG 1072
Db 803 TTATAGGATTAGAAATGTTTAAACAACATTAACAACAGCAACATCCGGAGATATGTAG 862
Qy 1073 GGTGCTGCTCGGTGATGTCAGAGAGGATGACATTTGAAAGAGGATGCTGCTGCTGCTGCT 1132
Db 863 GTATTTTAAAGAAATATTCAAAAAAGATATAAAAAAGAGGTATGATTTTGAACAAC 922
Qy 1133 CTGGCTCTATCACACCGGACACCAAGTTGAGGCTGTTGATGATGCTGCTGCTGCTGCTGCT 1192
Db 923 CTAATAAATTAAGATATATAAGTCTTTTATAGCTCAAAACATATATTTTAACTAAAG 982
Qy 1193 AGGGTGGCGCACCTCACCTTTCTTCTCCCTGCTTACCGCCACAGTTCTACATCGCGACAA 1252
Db 983 AAGTGGTCTCATAAACCTTTTATATTTGATATTAACCTCAATTTTTTATTCGTACAG 1042
Qy 1253 CTGATGTGACAGGAGTGTGACTAGCATTTATGAATGACAAGGATGAGGAGCGAGATGT 1312

Db 1043 TAGATGTTACTG--GAGAAATTAATAATATATATTTAAATGAAATGTACAAAAAGTAG 1099
 QY 1313 GCATCCCTGGTGACCGTATCAAAATGATTTTCAGCTCATCCAGCCCTGTCCTTGTGAGC 1372
 Db 1100 CTATCCCTGGAGATAAAATACATATACATATTTGAATTAATAACATATATATAGTGTGACAT 1159
 QY 1373 AGGATATGAGGTTTCTATCCGTGAGGTTGTAAGACCGTTGGTGCCGGTGTGATCAACA 1432
 Db 1160 TAAATATCAAAATTTCTATTATAGAGAGGAGGAAAAACAATAAGGACGAGGTATTATATAACAG 1219
 QY 1433 AAATCA 1438
 Db 1220 AAATAA 1225

RESULT 9

US-08-920-812-10/c
 ; Sequence 10, Application US/08920812
 ; Patent No. 5763188
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/920,812
 ; FILING DATE: 29-AUG-1997
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/362,577
 ; FILING DATE: 27-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 19036/32420
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3719 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecalis
 ; STRAIN: Clinical Isolate S2-3
 US-08-920-812-10

Query Match 10.8%; Score 182.2; DB 1; Length 3719;
 Best Local Similarity 58.2%; Pred. No. 2.8e-33;
 Matches 386; Conservative 0; Mismatches 258; Indels 19; Gaps 3;

QY 787 GCGGACGAGTGGTGGCTGACTACATCTTCTCGTTGGTTGATAAGTGATTCCTATATT 846
 Db 3708 GCGGACGAGTCTTATGAGAAAAAATCTAGAAATTAATGGCTGCAGTTGACGAATATATC 3649

QY 847 CCAGTCCCGCAGGAGCAGACTGACCTCCCGTCTTCTGCTCGCTGTTGAAGATGTCCTCTCC 906
 Db 3648 CCAACTCCAGAGAGTGTACTGACAAACCATTCATGATGCCAGTGAAGACGATTTCTCA 3589
 QY 907 ATCACCCTGGTGGTACAGTTGCACTGCGCGGTATAGAGCGTGGACCGCTCAAGATTTGTT 966
 Db 3588 ATCACTGGAGCTGGTACTGTTGCTACAGGCGGTGTTGAACGTTGGAAGTTCGCGCTTGT 3529
 QY 967 GACACATCGATATCGTGGAAATCCGGGAC---ACCCGGAACGTGACGCTCACTGGTGT 1023
 Db 3528 GACGAAGTTGAAATCGTTGGTATTAAGAGCAAGCAACATCTAAACAACACTGTTACAGGTGT 3469
 QY 1024 GAGATGTTCCAGAAACCATGATGCTATGGCCGAGACAAATGTTGGCTGCTGCTC 1083
 Db 3468 GAAATGTTCCGTAATATTAGACTAGCTGAAGAGGCGGACACATCGTGTCTTTATTA 3409
 QY 1084 CGTGTATGCAGAAAGATGACATTTGAAAGAGGCGATGCTGGCAAAGCCTGGCTCTATC 1143
 Db 3408 CGTGGTGTTCACGCTGAAGATATCGAACGTTGGACAAGTATTAGCTAAACGAGCTACAATC 3349
 QY 1144 ACACCCGACACCAAGTTTGAGGCTGTTGATGTATGCTTTAAGAAGAGAGGCTGCGCA 1203
 Db 3348 ACTCCACACACAAAATTCAAAGCTGAAGTATACGTATTATCAAAGAGAGAGGCGGACGT 3289
 QY 1204 CACTCACTCTTCTTCCCTGCTTACCCTCCACAGTTCTACATGCGGACAACTGATGTGACA 1263
 Db 3288 CACACTCA-TTCTTCACTAATCTGCTCTCAATCTACTTCCGTACACAGACGTTACT 3230
 QY 1264 GCGAGTGTGACTAGCATTTAATGACAAAGGATGAGGAGGCGGAAGATGTGCAATGCTGT 1323
 Db 3229 GGTGTGT-----AGAAATGCCACAGGTAAGTAAATGTTGCTGCT 3185
 QY 1324 GACCGTATCAAAATGATTGTTGAGCTCAATCCAGCTGTTGCTGTCAGCAGGATGAGG 1383
 Db 3184 GATAAGCTTGTATGAGCAGTTGAATTAATCAACCANTCGCTATCGAAGACGAACTCT 3125
 QY 1384 TTTGCTATCCGTGAGGTTGGTAAGACCGTTGGTGGCGGTGTCATCAACAAAATCATTTAG 1443
 Db 3124 TTCTCTATTCTGTAAGGCGGACGTAAGTTAGGTTACGGCGTGTGTTACTGAAATCGTTAAA 3065
 QY 1444 TAA 1446
 Db 3064 TAA 3062

RESULT 10

US-08-920-827-10/c
 ; Sequence 10, Application US/08920827
 ; Patent No. 5770375
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/920,827
 ; FILING DATE: 29-AUG-1997
 ; CLASSIFICATION: 435

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,577
/ FILING DATE: 27-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 19036/32420
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3719 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecalis
/ STRAIN: Clinical Isolate S2-3
/ US-08-920-827-10

Query Match 10.8%; Score 182.2; DB 1; Length 3719;
Best Local Similarity 58.2%; Pred. No. 2.8e-33;
Matches 386; Conservative 0; Mismatches 258; Indels 19; Gaps 3;

QY 787 GCGGACGATGAGTGGTGGCTACATCTCTCGTTGGTTGATAAAGTGCATTCCTATATT 846
Db 3708 GCGGACGATCTTATGAAGAAAATCTTAGAATTAATGGCTGCGAGTTGACGAATATATC 3649

QY 847 CCAGTCCCGCAGAGGAGACTGACCTCCCGTTCTTGTGCTGGTGTGAAGATGCTTCTCC 906
Db 3648 CCAACTCCAGAGAGCTGATACATGACAAACCATTCATGATGCCAGTGAAGACGTATCTCA 3589

QY 907 ATCACCAGTGTGGTACATGTCGCTGATGAGCGGTGACCGTCAAGATTTGGT 966
Db 3588 ATCACTGGACGTGGTACTGTTGCTACAGCCGCTGTTGAACGCTGGTGAAGTTCGGCTGGT 3529

QY 967 GACACAGTCGATCTGCGGATCCGGAC---ACCCGGAAGTGCACGCTACGCTGTT 1023
Db 3528 GAGGAAGTGAATCGTTGGTATTAAGACGAACATCTAAACAACTGTTACAGGTTT 3469

QY 1024 GAGATGTTCCAGAGACCATGATGCCATGGCGGAGACAATGTTGGGCTGCTGTC 1083
Db 3468 GAAATGTTCCGTAATTTATGACTACGCTGAAGCAGCGACACATCGTGCCTTATTA 3409

QY 1084 CGTGGTATCCAGAGGATGACATTCGAGAGCGATGGTGGCAAGCCCTGGCTATC 1143
Db 3408 CGTGGTGTGACGTGAAGATATCGAAGCTGGACAAGTATTAGCTTAAACACGCTACATC 3349

QY 1144 ACACCGCACACCAAGTTTGGGCTGTTGATGCTTAAAGAGGAGAGGTTGGCGA 1203
Db 3348 ACTCCACACACAAATTCAGAGCTGAAGTATACGATTATCAAGAGAGGCGGAGCT 3289

QY 1204 CACTCACCTTCTCCCTGGTTACCGCCAGCTTCTACATCGCGACAACATGATGACA 1263
Db 3288 CACACTCA-TTCTTCACTAACTATCGTCTCAATTTCTACTTCCGTACACAGAGCTTACT 3230

QY 1264 GGGAGTGTGACTACGATTATGATGACAAAGATGAGGAGGCGAGATGTCATGCCCTGGT 1323
Db 3229 GGTGTTGT-----AGAAATTGCCAAGAGGCTACTGAAATGGTAAATGCCCTGGT 3185

QY 1324 GACCGTATCAAAATGATTTCAGCTCATCCAGCTGTTGCTGTTGAGCAGGATATGAGG 1383
Db 3184 GATACGTTGCTATGAGAGCTTGAATTAATCACCATACTGCTATCGAAGAGCGGAACCTGT 3125

QY 1384 TTGCTATCCGAGGGGTGAAGACCGTTGGTCCCGGTGTCATCAACAAATCATTTGAG 1443
Db 3124 TTCTTAATTCGTGAAGCGGACGCTACTGTAGTTTCAGCGCTTGTACTGAAATCGTTAA 3065

QY 1444 TAA 1446
|||

Db 3064 TAA 3062

RESULT 11
US-08-921-177-10/c
: Sequence 10, Application US/08921177
: Patent No. 5798211
: GENERAL INFORMATION:
: APPLICANT: Ohno, Tsuneya
: APPLICANT: Matsuhisa, Akio
: APPLICANT: Uehara, Hirotugu
: APPLICANT: Eda, Soji
: TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, version #1.25
: CURRENT APPLICATION DATA: US/08/921,177
: FILING DATE: 29-AUG-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/362,577
: FILING DATE: 27-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Rin-Laures, Li-Hsien
: REGISTRATION NUMBER: 33,547
: REFERENCE/DOCKET NUMBER: 19036/32420
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3719 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: Genomic DNA
: ORIGINAL SOURCE:
: ORGANISM: Enterococcus faecalis
: STRAIN: Clinical Isolate S2-3
: US-08-921-177-10

Query Match 10.8%; Score 182.2; DB 1; Length 3719;
Best Local Similarity 58.2%; Pred. No. 2.8e-33;
Matches 386; Conservative 0; Mismatches 258; Indels 19; Gaps 3;
```


STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920.828
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3719 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecalis
STRAIN: Clinical Isolate S2-3
US-08-920-828-10

Query Match 10.8%; Score 182.2; DB 2; Length 3719;
Best Local Similarity 58.2%; Pred. No. 2.8e-33;
Matches 386; Conservative 0; Mismatches 258; Indels 19; Gaps 3;

QY 787 GCGCAGCATGAGTGGTGGCTACATCTTCTGTTGTTGATAAAGTGGATTCCTATATT 846
|||||
DB 3708 GCGCAGCATGTTATGAAGAAAATCTTAGAATTAATGCTGCAGTTGACGAATATATC 3649
|||
QY 847 CCACTCCCGCAGAGCCAGACTGACCTCCGTTCTTGTCTGCTGTGAAGATGCTTCTCC 906
|||
DB 3648 CCAACTCCAGAACGTGATACTGACAAACCATTCATGATGCCAGTGAAGACGTATTCTCA 3589
|||||
QY 907 ATCACCCTGCTGTCAGTTCACCTGCCCTGATGAGCGTGGCCACCGTCAAGATTGGT 966
|||||
DB 3588 ATCACTGGACGTGCTACTGTGTGTACAGCCGCTGTTGAACGTGGTGAAGTTCGCGTTGGT 3529
|||||
QY 967 GACACATGATATGTCGGAATCCGGGAC---ACCCGGAACGTGCAGCTGCTGCTTT 1023
|||
DB 3528 GACCAAGTTGAATCGTGTGTTAATGAAGAGAAACATCTAAACAACTGTACAGGTGT 3469
|||||
QY 1024 GAGATGTTCCAGAGACCATGATGATGCCCGGAGACAAATGTTGGCTGCTGCTC 1083
|||||
DB 3468 GAAATGTTCCGTAATATTATTAGACTAGCTGAAGAGGCGACAAATCGTGTGCTTTATTA 3409
|||||
QY 1084 CGTGGTATGCAAGGATGACATTGAAGAGGCATGGTCTGCAAGCGCTGCTCTATTC 1143
|||||
DB 3408 CGTGGTGTCCAGTGAAGATATCGAACGTGGACAGATATTAGCTAAACCACTACAATC 3349
|||||
QY 1144 ACACCGCACCAAGTTTGAAGCTGTTGTGTATGTCCTTAAAGAAAGAGAGGTGGCCCA 1203
|||||
DB 3348 ACTCCACACACAAAATTCAAAGCTGAAGTATAGTATTATCAAAAGAAAGGCGGACGT 3289
|||||
QY 1204 CACTCACCTTCTTCCCTGGTTACCCGCCACAGTCTTACATCGCGGACAACTGATGTGACA 1263
|||||
DB 3288 CACACTCA-TTCTTCACTAACTATCTGCTCTCAATTCTCTTCCGTACACACAGACGTACT 3230
|||||

QY 1264 GGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGGAAGATGTCATCCCTGGT 1323
|||
DB 3229 GGTGTTGT-----AGAAATGCGAGAAGTACTGAATGGTAAATGCTCTGGT 3185
|||
QY 1324 GACCGTATCAAAATGATTGTTACGCTCATCCAGCTGTGCTTGTGAGCAGGGGTATGAGG 1383
|||
DB 3184 GATAACGTTGCTATGGACGTTGAATTAATTCACCCAATCGCTATCGAAGACGGAACTCGT 3125
|||
QY 1384 TTTGCTATCCGTGAGGTGTTAAGACCGTGGTCCCGGTGTCATCAACAAAATCATTTGAG 1443
|||
DB 3124 TTCTCTATTCGTGAAGCGGACGTTACTGTAGTTTCAGCGCTGTTTACTGAATCGTTAAA 3065
|||
QY 1444 TAA 1446
|||
DB 3064 TAA 3062

RESULT 14

US-08-936-165A-87
; Sequence 87, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-87

Query Match 10.7%; Score 180.2; DB 4; Length 656;
Best Local Similarity 62.4%; Pred. No. 4.3e-33;

Matches 295; Conservative 0; Mismatches 177; Indels 1; Gaps 1;
QY 286 GGAAGACCACTCTCACCGCGGCTCACCATGGTGTCTCCCTCCGTTCGGTGGCAGCGG 345
Db 3 GGTAAACAACATTAACAGCAGCAATCGCTACTGTATTAGCAAAAATGGTACTCAGTT 62
QY 346 CCTAAGAGTACGACGAGATCGAGCGCGCCCGCGAGGAGCGCGCGGTATACCATC 405
Db 63 GCACAATCATATGATGATTTGACACGCTCCAGAGAAAAGAACGCTGTATCACAATC 122
QY 406 AACACCGCCACGCTCGAGTACGACGAGACCGCGCCACTACGACACGCTCGACTGCCCC 465
Db 123 AATACTTCTACATTTAGTACCANACTGACAAAGTCACTANGCTCAGCTCGNCTGCCA 182
QY 466 GG-CCACCGCGGACTATGCAAGAAATATGATCACCGCGCTGCGGAGATGACGCGTGCAT 524
Db 183 GNAATCCGCTGACTACGTTAAACACATGATCACTGGTGTCTCAAAATGGACGCGGTAT 242
QY 525 CTTGCTGTATCCGGTCCGAGCGGCCCATGCGCGAGACCAAGAGCACATCTCTCTCGC 584
Db 243 CTTAGTAGTATCTGCTGACGCTCCAAATGCCACAACTCGTGAACACATCTTTTATC 302
QY 585 CAAGCAAGTCGGTGTCCCAAGATCGTTGCTCTCTCAACAAGAGACATGTCGACGA 644
Db 303 ACGTAACGTTGGTGTACGAGCATTAGTAGTATTCTTAAACAAAGTTGACATGGTTGACGA 362
QY 645 CGAGGAGTGTCTGAGCTCGTGCAGCTCGAGGTCGGGAGCTGCTCAGCAACTACGAGTA 704
Db 363 TGAAGAATTATTAGAATTAGTAGAAATGGAAGTTGCTGACTTATTAAGCGAATATGACTT 422
QY 705 CGACGCGGACGAGTACCAATCGTCTGCTGCTCCGCTCCGCTCAAGCGCTCGAGG 757
Db 423 CCCAGGTGACGATGTACCTGTAATCGCTGGTTCAGCATTAAGAGCTTTAGAG 475

RESULT 15
US-09-556-877-70/c
; Sequence 70, Application US/095556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556.877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 70
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-556-877-70
Query Match 6.7%; Score 112.6; DB 4; Length 323;
Best Local Similarity 60.7%; Pred. No. 1.9e-17;
Matches 184; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 879 CTTGCTCGCTGTTGAAGATGTCTTCCCATCACCGGTCGTGTACAGTGTGCCACTGGCGG 938
Db 323 CTTAATGCCCATGAGAGCGTGTCTCTATCTCCGACGAGGAAGTGTAGTAACCTGGACG 264
QY 939 TATAGAGCTGGCAGCGTCAAGATGGTGACACAGTCTGATATCGTCGGGAATCCGGGACAC 998
Db 263 TATTGAGCGTGAATTTGTTAAAGTTTCCCGATAAAGTTCAAGTTGGTGGTCTTTAGAGATAC 204
QY 999 CCGGAAGTGCACGCTGCTGGTGTTCAGATGTTCCAGAGACCATGATGATGCCATGCC 1058
Db 203 TAAAGAAACGATGTTTACTGGGGTTGAAATGTTTCAAGAAAGAACTCCCAAGAGTCGTGC 144

QY 1059 CGGAGACAATGTTGGCTGCTGCTCCGTGGTATGCAAGGATGACATTTGAAAGAGGCAAT 1118
Db 143 AGGAGAGAAATGTTGGATTGCTCCTCAGAGGTATTGGTAAGAACGATGTGGAAGAGGAAT 84
QY 1119 GGTGCTGGCAAGCCCTGCTCTATCACACCGGACACCAAGTTTGAGGCTGTTGTGTATGT 1178
Db 83 GGTGCTGTTGCTTGCCAAACAGAGTGTAAACCTCATACAGGTTTAAAGTGTGCTGTTAGCT 24
QY 1179 GCT 1181
Db 23 TCT 21

Search completed: May 1, 2003, 11:17:53
Job time : 3545 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 15:58:02 ; Search time 273 Seconds
(without alignments)
7275.843 Million cell updates/sec

Title: US-09-810-764A-6

Perfect score: 1681

Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 745064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.4	31.4	1194	10	US-09-815-242-7897
2	521.6	31.0	1194	10	Sequence 7897, Ap
3	508.8	30.3	1185	10	Sequence 7904, Ap
4	508.8	30.3	1185	10	Sequence 86, Appl
5	505.6	30.1	1230	10	Sequence 6265, Ap
6	503.6	30.0	1185	10	Sequence 9866, Ap
7	450.2	26.8	3309400	9	Sequence 6334, Ap
8	446.4	26.6	1188	9	Sequence 1, Appli
9	436	25.9	1200	10	Sequence 3437, Ap
10	436	25.9	1448	8	Sequence 7318, Ap
11	420.4	25.0	1185	10	Sequence 3, Appli
12	420.4	25.0	1185	10	Sequence 6962, Ap
13	413.8	24.6	1057	10	Sequence 6972, Ap
14	396.4	23.6	1197	10	Sequence 1309, Ap
15	394	23.4	15598	7	Sequence 9073, Ap
16	393.2	23.4	1185	10	Sequence 82, Appl
17	390.8	23.2	1182	10	Sequence 8344, Ap
18	378.6	22.5	1188	10	Sequence 4519, Ap
19	376.2	22.4	1185	10	Sequence 6385, Ap
					Sequence 3832, Ap

Sequence 575, App
Sequence 166, App
Sequence 164, App
Sequence 1, Appl
Sequence 154, App
Sequence 526, App
Sequence 88, Appl
Sequence 151, App
Sequence 171, App
Sequence 159, App
Sequence 1, Appl
Sequence 168, App
Sequence 456, App
Sequence 52, Appl
Sequence 155, App
Sequence 158, App
Sequence 147, App
Sequence 162, App
Sequence 160, App
Sequence 156, App
Sequence 152, App
Sequence 148, App
Sequence 18, Appl
Sequence 165, App
Sequence 157, App
Sequence 125, App

ALIGNMENTS

RESULT 1
US-09-815-242-7897
; Sequence 7897, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 7897
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1194)
US-09-815-242-7897

Query Match 31.4%; Score 528.4; DB 10; Length 1194;
Best Local Similarity 66.1%; Pred. No. 5.3e-133;
Matches 813; Conservative 0; Mismatches 381; Indels 36; Gaps 2;

QY 217 GCGGCGAGGGCAAGTTCGAGCGGACCAACACACACGTCACATAGGACGACCATCGGCAT 276
Db 1 GTGGCTAAGGAAAAATTCGACGTAAACAAACCGCACGTCAACGTCGCGACCATCGGTAC 60

QY 277 GTCGACACCGGAAAGACCACTCTACCGCGCGCTGACCATGGTGTCTCGCTCGCTCGGT 336
Db 61 GTTGACCATGGCAAGACCACTCTGACCGCTGCATGACCAAGGTCTGCTCCGATACCTGG 120

QY 337 GCGACGCGCCCTAAGAAGTACGACGATCGACGCGCCCGCGAGGAGCGCGCGGGT 396
Db 121 GGTGGTTCGGCTCGGTCTTCGATCGATCGACACGCGCGGAGAAAGCGCGCGGT 180

QY 397 ATCACCATCAACACCGCCACCGTCTGAGTACGAGACCGGACCGCCACTAGCACACGTC 456
Db 181 ATCACCATCAACACCGCCACCGTCTGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 240

QY 457 GACTCGCCCGGCGGACGCTATGTCGAGTATGATCAGCGCGCTGCGCAGATGAC 516
Db 241 GACTCGCCCGGCGGACGCTATGTCGAGTATGATCAGCGCGCTGCGCAGATGAC 300

QY 517 GGTGCGATCTCGTGTATCCGGTATCCGGTCCGACGCGGCGGCGGCGGCGGCGGCGG 576
Db 301 GGTGCGATCTCGTGTATCCGGTATCCGGTCCGACGCGGCGGCGGCGGCGGCGGCGG 360

QY 577 CTCCTCGCAACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db 361 CTCCTCGCAACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 637 CTCGACGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db 421 CTCGACGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 697 TACGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 756
Db 481 TACGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 757 GCTCTCATGGTCAACCCCTGCTTGAAGCGCGGCGGACGATGAGTGGTGCAGTATCTTC 816
Db 541 G-----GCAAGGATGACAAACGAGTTCGCGGTAAGCGCGCGTCCAG 579

QY 817 TCGTTGGTTGATAAGTGGATTCCTATATCCAGTCCGCGAGAGGAGGAGGAGGAGGAGG 876
Db 580 TCGTTGGTTGATAAGTGGATTCCTATATCCAGTCCGCGAGGAGGAGGAGGAGGAGGAG 639

QY 877 TTCCTGCTGCTGTTGAAGATGCTTCTTCATCAGCGGTCGTTGATGATGATGATGATG 936
Db 640 TTCCTGCTGCTGTTGAAGATGCTTCTTCATCAGCGGTCGTTGATGATGATGATGATG 699

QY 937 GCTATAGAGCTGGCACCGTCAAGATGGTGACACAGTTCGATATCTCGGAATCCGGGAC 996
Db 700 GCTATAGAGCTGGCACCGTCAAGATGGTGACACAGTTCGATATCTCGGAATCCGGGAC 759

QY 997 ACCCGGAACTGCAGCGTCACTGTTGATGATGTTCCAGAGACCATGATGATGATGATG 1056
Db 760 ACCCGGAACTGCAGCGTCACTGTTGATGATGTTCCAGAGACCATGATGATGATGATG 819

QY 1057 GCGGAGACATGTTGGCTCTGCTCGGTGATGACAGAGGATGATGATGATGATGATGATG 1116
Db 820 GCTGTTGAGACGTTGGTATCTCTGCTGGGACCAAGCGTGAAGACGTTAGAGCGTGGC 879

QY 1117 ATGTTGCTGGCAAGCTGGCTCTATCACACCGCACACCAAGTTGAGGCTGTTGTTAT 1176
Db 880 CAGGTACTGGCCAAAGCGGGACCATCAAGCGGACACCAAGTTCGAGTTCGAGTGTAC 939

QY 1177 GTGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1236
Db 940 GTGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 999

QY 1237 TTCTACATCGGACAACTGATGTACAGGAGGTGTGACTACGATTAATGATGACAAAGAT 1296
Db 1000 TTCTACTTCGTTACCAACCGATGTGACCGG-----TAACTGCGAAGTCCGCG 1044

QY 1297 GAGGAGCGAAGATGTCATGCTGCTGAGCCGTATCAAAATGATTTTCAGCTCATCCAG 1356
Db 1045 GAAGGCGTGAAGATGTAATGCGCGGCGACAAATCAAGATGTTGTCACCTGATCGCT 1104

QY 1357 CTTGTTGTTGTCAGCAGGATGAGTGTGCTATCCGTGAGGTGTTAAGACCGTGTGT 1416
Db 1105 CCATCGCATGGAAGATGGCTGCGCTTCGCGATCCGGAAGCGCGCTACCGTTGGC 1164

QY 1417 GCGCGTGTCAACAACAAATCATTTAGTAA 1446
Db 1165 GCGCGGTGTTGCCAAGATCATCGAGTAA 1194

RESULT 2
US-09-815-242-7904
; Sequence 7904, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7904
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1194)
US-09-815-242-7904

Query Match 31.0%; Score 521.6; DB 10; Length 1194;
Best Local Similarity 65.8%; Pred. No. 3.6e-131;
Matches 808; Conservative 0; Mismatches 384; Indels 36; Gaps 2;

QY 219 GCGGAGGGCAAGTTCGAGCGCACCAACACACGTCACATAGGACGACCATCGGCATGT 278
Db 3 GCGTAAAGAAAATTTGAACGGAACACGCGCACGTCACGTCGCGACCATCGGTACGT 62

QY 279 CGACACGGAAGAACACATCTCACCGCGCGCTCACCATGCTGCTCGCTCGGTGG 338
Db 63 TGACCATGGAAGAACACATCTCACCGCGCTCACCATGCTGCTCGGTGG 122

QY 339 CAGCGCGCTTAAGAGTACGACGAGATCGACCGCGCGCGCGCGCGCGCGGTAT 398

Db 123 TGGTTCGGCTCGTCTTCGATCAGATCGACAGCGCCGGAAGAAAGCCCGGGTAT 182
 Qy 399 CACCATCAACACCCGACCGTTCGATGACGAGACCGCCGACCTACGACACGTCGA 458
 Db 183 CACCATCAACACCCGACCGTTCGATGACGAGACCGCCGACCTACGACACGTCGA 242
 Qy 459 CTGCGCCGGCCGACCGGCTATGTCGAAGATATGTCACCGGGGCTGCGAGATGACGG 518
 Db 243 CTGCGCCGGTCAACCGGCTACGTCGAAGACATGATCAGCGGTGCTGCCAGATGACGG 302
 Qy 519 TGGCATCTCTGCTATTCGGTTCGCGGACCGGCCCCATCCGACGACCAAGAGACATCCT 578
 Db 303 CGCGATCTCTGCTTCGGTTCGCGGACCGGCCCCATCCGACGACCGGACATCCT 362
 Qy 579 CTTGCGCAAGCAAGTCTGTTCCCAAGATCGTTTCTTCTTCTCAACAGAGACATGTT 638
 Db 363 CTTGCGCGGAGTAGCGTTCCTTACATCGTCTGTTCTTCTTCAACAGAGACATGTT 422
 Qy 639 CGAGGAGGAGTCTGCTCGAGTCTGCTGAGTCTGAGGTCGCGGAGTCTGCTCAGCACTA 698
 Db 423 CGAGGAGGAGTCTGCTGAGTCTGCTGAGATGGAATGCTGCGATCTGCTGACACCTA 482
 Qy 699 CGAGTACGAGGCGACGACGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
 Db 483 CGACTTCGCGGCGACGACATCTCGATCATCTCGGTTCCGCGCTGATGCGCTGGAAG - 541
 Qy 759 TCTCATGTCACACCTCCCTTGAAGCGCGGACGATGAGTGGTGCATACATCTCTC 818
 Db 542 -----GCAAGGATGACAACCGCATCGGCTGAAGCGCCGTCAGAA 581
 Qy 819 GTTGGTGTAAAGTGGATCTTATATTCAGTCCCGGACGACGACGACTGACCTCCGCT 878
 Db 582 GCTGTAGAGACCTGACTCTTACATTCGAGCGCGTTCGTCGCTACGACGCGT 641
 Qy 879 CTTGCTCGTGTGAAGATCTTCTCCATCACCCTGCTGCTGCTGCTGCTGCTGCTGCTG 938
 Db 642 CTTGATCCGATCGAAGACGTTCTCGATCTCCGCGCGGTACCGTGTGAACCGCTG 701
 Qy 939 TATAGACGCTGGACCGCTCAAGATTTGTTGACAGTCTGATCTGCGAATCCGGACAC 998
 Db 702 TGTAGAGCGGACATCATCAAGTCCAGGAAGTGGAAATCTGCGCATCAAGCGGAC 761
 Qy 999 CCGGAATCGACGCTACTGCTGTTGAGATGTTCCAGAGACCATGATGATGCCATGCC 1058
 Db 762 CACCAAGACTACCTGACCGCGTTGAAATGTTCCGCAAGCTCTCGACGAAGTCTGTC 821
 Qy 1059 CGGAGCAATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
 Db 822 TGTGGAACGTTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
 Qy 1119 GGTGCTGGCAAGGCTGCTGCTATCACCGGACACACCAAGTTTGAAGGCTGTTGATGT 1178
 Db 882 GTTCTGGCCAGCGCGGACCATCATAGCGGACACCAAGTTTGAAGGCTGTTGATGT 941
 Qy 1179 GCTTAAAGAGAGAGGTTGCGGACACTCACCTTTCTTCCCTGTTTACCGCCACAGTT 1238
 Db 942 GCTGTCAAGGAAGAGTGTGCTCACACCCGCTTCTTCAAGGGTACCGTCCGAGTT 1001
 Qy 1239 CTACATCGGACAACTGATGTGACAGGAGTGTGACTACGATTTGAATGACAAGGATGA 1298
 Db 1002 CTACTTCGTTACCAACGACGTGACCGG - -----TAACTGGGAAGTCCGCGA 1046
 Qy 1299 GGAGGCGAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
 Db 1047 AGCGGTAGAGATGTTATGCGCGGCGACAAACATCAAGATGTTGTCACCGTACGCTCC 1106
 Qy 1359 TGTGCTTGTGACAGGATGATGAGTTGCTATCCGTTGAGGGTGTGAAGACCGCTGCTGTC 1418
 Db 1107 GATCGCCATGAAGATGCGCTGCTGCGGATCCGCGAGGCGCGGTACCGCTTGGCGC 1166
 Qy 1419 CGGTGTCATCAACAAATCATTTGAGTAA 1446

Db 1167 CGCGCTGGTTCGAAGATCATCGAATAA 1194
 RESULT 3
 ; Sequence 86, Application US/09912020
 ; Patent No. US20020045592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Trawick, John
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Froelich, Jamie M.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
 ; FILE OF INVENTION: ESCHERICHIA COLI
 ; FILE REFERENCE: ELITRA.001DV1
 ; CURRENT APPLICATION NUMBER: US/09/912,020
 ; PRIOR FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: 09/492,709
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: 60/117,405
 ; PRIOR FILING DATE: 1999-01-27
 ; NUMBER OF SEQ ID NOS: 485
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 86
 ; LENGTH: 1185
 ; TYPE: DNA
 ; ORGANISM: E. Coli
 ; US-09-912-020-86
 Query Match 30.3%; Score 508.8; DB 10; Length 1185;
 Best Local Similarity 65.3%; Pred. No. 1.1e-127;
 Matches 803; Conservative 0; Mismatches 382; Indels 45; Gaps 2;
 Qy 217 GCGGCGAGGGGGAAGTTTCGAGCGCACCAACACACGTCACATAGGACACCATCGGCCAT 276
 Db 1 GTGCTTAAGAAATTTGAACGTACAAACCGCAGGTTAAGTTGGTACTATCGGCCAC 60
 Qy 277 GTCGACACGGAAGAACCACTCTCACCGCGGCTACACATGGTCTCGCCTCCGCTCGGT 336
 Db 61 GTTGACACCGTAAACTACTCTGACCGCTGCAATCACCACCGTACTGGCTAAACCTAC 120
 Qy 337 GCGAGGCGCCTTAAGAGTACACGAGATCGACGCGCGCCCGGAGGAGGCGCGCGGT 396
 Db 121 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Qy 397 ATCACCATCAACACCGCCACCGTTCGAGTACGAGACCGGACCGCCGCTACGACACGTC 456
 Db 181 ATCACCATCAACACTTCTACGTTGAATACGACACCGCCGACCGCTACTACGACACG 240
 Qy 457 GACTGCCCGCGGACCGGCTATGTCGAAGATATGTCACCGGCGCTGCGGAGATGGAC 516
 Db 241 GACTGCCCGGCGACCGGCTATGTTAAACATGATCACCAGTGTGCTCAGATGGAC 300
 Qy 517 GGTGCAATCTCTGCTATTCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 576
 Db 301 GCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Qy 577 CTCTTCGCCCAAGTGTGCTTCCAGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTT 636
 Db 361 CTGCTGGGCTGCTGAGTAGGCGTTCGCTACATCATCTGTTCTTCTTCTTCTTCTTCT 420
 Qy 637 GTCGACGACGAGGAGTGTCTGAGCTCGTTCGAGCTCGAGGTCGCGGAGCTGTCTCAGAAC 696
 Db 421 GTTGTATGACGAAGAGCTGCTGGAAGTGTGGAAGTGTGGAAGTGTGGAAGTGTGGAAG 480
 Qy 697 TACGATGACGAGGCGGACGACGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 756
 Db 481 TACGACTTCCCGGCGGACGACACTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539

QY 757 GCTCTCATGGTCAACCCCTGCTTGAAGCGGGCGGACGATGAGTGGGTCGACTACATCTTC 816
Db 540 -----AGGCGACGAGTGGGAAGCGAAAATCCTG 570
QY 817 TCGTTGGTTGATAAAGTGGATTCCTATATTTCCAGTCCCGCAGAGCGAGACTGACCTCCG 876
Db 571 GAACCTGGCTGCTTCCGTGGATTCCTATATTTCCGGAACGAGCGGTGGATTTGACAAGCCG 630
QY 877 TCTTGTCTGCTGTGTAAGATGCTTCTTCCATCATCCGGTCTGTGTACAGTTGCCACTGGC 936
Db 631 TTCCTGCTGCCGATCGAAGAGCTATTCTCCATCTCCGGTCTGTGTACCTGTTGTACCGGT 690
QY 937 CGTATAGAGCGTGGACCGCTCAAGATGGTGACACATCGATATPCGTCGAAATCCGGGAC 996
Db 691 CGTGTAGAAGCGGTATCATCAAGTGGTGAAGAAGTTGAATCGTTGGTATCAAAAG 750
QY 997 ACCCGGAACGCGGTCACTGGTGTTCAGATGTTCCAGAGACCATGGATGATGCCATG 1056
Db 751 ACTCAGAAGTCTACTGTACTGGGCTGGAATGTTCCGCAACTGCTGGAGCGAAGCCGT 810
QY 1057 GCCGAGACAATGTTGGGCTGCTGCTCGGTGATGACAGAGGATGACATTGAAAGAGC 1116
Db 811 GCTGTGAGAAGCTAGTGTCTGCTGCTGTATCAACGCTGAAGAAATCGAAGCTGT 870
QY 1117 ATGGTCTGGCAAGCTGGCTCTATCACACCGCACACCAAGTTTGAGGCTGTGTGTAT 1176
Db 871 CAGGTACTGGCTAAGCGGGGACCATCAAGCCGACACCAAGTTTCAAGTCTGAAGTATC 930
QY 1177 GTGCTTAAGAAGAGGGTGGCGACACTCCTTCTTCCCTGCTTACCGCCACAG 1236
Db 931 ATTCGTCAAAGATGAAGCGCGCTCATCTACTCCGTTCTTCAAAGCTTACCGTCCGAC 990
QY 1237 TTCTACATCGCGACAACATGTCACAGGGAGTGTGACTACGATTTATGAATGACAAGGAT 1296
Db 991 TTCCTACTTCCGTACTGACGTGACTGG-----TACCATCGAAGTCCCG 1035
QY 1297 GAGGAGCGCAAGATGTCATCCCTGGTGGACCGTATCAAAATGATTTTACGCTCATCCAG 1356
Db 1036 GAAGCGTAGAGATGGTAAATCCCGGGGACCAACATCAAAATGGTTGTACCTCGATCCAC 1095
QY 1357 CTTGTGCTGTGACGAGGATGAGGTTTGTCTATCCGTGAGGTTGGTAAAGACCGTTGGT 1416
Db 1096 CCGATCGGATGGACGAGGCTCGGTTTCGCAATCCGTGAAGCGCGCGTACCGTTGGC 1155
QY 1417 GCCGCTCATCAACAAATCATTTAGTAA 1446
Db 1156 GCGGCGGTGTTGCTAAAGTTCTGGGCTAA 1185

RESULT 4

US-09-815-242-6265

; Sequence 6265, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6265
; LENGTH: 1185

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)....(1185)

US-09-815-242-6265

Query Match 30.3%; Score 508.8; DB 10; Length 1185;

Best Local Similarity 65.3%; Pred. No. 1.1e-127;

Matches 803; Conservative 0; Mismatches 382; Indels 45; Gaps 2;

QY 217 GCGGCGAGGGGCAAGTTCAGCGCACCAACACACAGTCAACATAGGACCATCGGCCAT 276
Db 1 GTGTCTAAAGAAAAATTTGAACGTACAAAACCGCAGCTTAACGTTGGTACTATCGGCCAC 60
QY 277 GTGACACCGGAAAGACCACTCTACCGCGCGGCTCACCATGGTGTGCTCCCTCCGT 336
Db 61 GTTGACCAAGGTAAACTACTCTGACCGTGCATCAACACCGCTACTGCTAAACCTAC 120
QY 337 GCGAGCGCCCTAAGAAGTACAGAGATCGAGCGGCGCCCGAGAGCGCGCGGT 396
Db 121 GCGGTGCTGCTGCTGATTCGACAGATCGATAACGCGCGGAGAAAAAGCTGTGT 180
QY 397 ATCACCATCAACCGCCACCGTCGAGTACGAGACGAGACCGCCGCTACACACACGTC 456
Db 181 ATCACCATCAACACTTCTCAGCTTGAATACGACACCGCCGCTACATCCGACACGTA 240
QY 457 GACTGCCCGGCGCACCGGACTATGTCAAGAATATGATCACCGGCTGCGGATGAGAC 516
Db 241 GACTGCCCGGCGCACCGGACTATGTAAACACATGATCACCGTCTGCTCAGATGGAC 300
QY 517 GGTGCCATCTGCTGCTATCCGTCGCGGCGCCCATCGCGACGACCAACAGACATC 576
Db 301 GCGCGATCTCTGTAGTTGCTGCGACTGACGCGCGGATCCGCGACACTCGTGAGCACATC 360
QY 577 CTCCTCGCCCAAGCAAGTCGGTGTCCCAAGATGCTTCTTCTTCAACAAGAGACATG 636
Db 361 CTGCTGGTCTGTCAGTAGGCTTCGCTACATCATCTGTTCTTGAACAAATCGACATG 420
QY 637 GTCGACGAGAGAGTCTCGAGCTCGTGGAGTCCGAGGTCGGGAGTGTCTACGACAC 696
Db 421 GTTGATGACGAAGAGCTGTGGAACCTGTTGAATGGAAGTTTCGTGAACCTTCTGTCTCAG 480
QY 697 TAGAGTACGAGCGGCGACGACGTACCAATCGTCGCTGGCTCCGCTCAAGCGCTCGAG 756
Db 481 TAGACTTCCGCGCGGCGACGACACTCGGATCGTTGCTGCTGTTGCTGCTGAAGCGGTGA- 539
QY 757 GCTCTCATGGTCAACCCCTGCCCTTGAAGCGCGGCGAGATGAGTGGTTCGACTACATCTTC 816
Db 540 -----AGGCGACGCGAGAGTGGGAAGCGAAATTCCTG 570
QY 817 TCGTTGGTTGATAAAGTGGATTCCTATATTTCCAGTCCCGCAGAGCGAGACTGACCTCCG 876
Db 571 GAACCTGGCTGGCTTCCGTGGATTCCTATATTTCCGGAACGAGAGCGTGCATTCACAAGCG 630
QY 877 TCTTCTGCTGCTTCTGAAGATGCTTCTCCATCACCGGTCGTTGGTACAGTTCCTCACTGCG 936
Db 631 TTCTGCTGCGGATCGAAGACGATTTCTCCATCTCCGTCGTTGGTACCTGTTGACCGGT 690
QY 937 CGTATAGAGCGTGGCACCGCTCAAGATTTGGTGACACAGTCGATATCGTCGGAATCCGGGAC 996

Db 691 CGTGTAGAACCGGTATCATCAAGTTGGTGAAGAAGTTGAAATCGTGTGATCAAAAGAG 750
QY 997 ACCCGAAGTGCACGGTCACTGTTGTAGATGTTCCAGAGACCATGATGATGCCATG 1056
Db 751 ACTCAGAAGTCTACCTGTTACTGCGTGAATGTTCCGCAAACTGCTGGACGAAGCCGT 810
QY 1057 GCGGAGACAAATGTTGGGCTGCTGCTCGGTGATGCAAGAAGGATGACATTGAAAGAGGC 1116
Db 811 GCTGGTGAGAAGCTAGTGTGTTCTGCTGCGTGGTATCAACAGTCAAGAAATCGAACGTGGT 870
QY 1117 ATGGTGTGGAAGCCGTGCTCTATCACACCGCACCAAGTTTGAGGCTGTGTTGTAT 1176
Db 871 CAGGTGTGCTTAAGCGCGCACCATCAAGCGCGCACCAAGTTTCAAGTGTGATC 930
QY 1177 GTGCTTAAGAAAGAGGGTGGCGCACACTCACTTCTTCCCTGTTACCGCCACAG 1236
Db 931 ATTCTGTCCAAAGATGAAGCGCGCTACTCCCTGTTCTTCAAGGCTACCGCTCCGCGAG 990
QY 1237 TTCTACATGCGGACAACTGATGTGACAGGGAGTGTGACTACGATTATGAATGACAAGGAT 1296
Db 991 TTCTACTTCTGCTACTACTGACGTGACTGG-----TACCATCGAAGTGGCG 1035
QY 1297 GAGAGCGGAAGATGTCATGCTGCTGCTGAGCCGTATCAAAATGATTTGTCAGCTCATCCAG 1356
Db 1036 GAAGCGGTAGAGATGTAATGTCGGGCGCACCAATCAAAATGTTGTTACCCCTGATCCAC 1095
QY 1357 CCGTGTGCTGTGAGCAGGGTATGAGTTTGTCTATCCGTGAGGTTGGTAAGACCGTTGGT 1416
Db 1096 CCGATCCGATGACAGCGGTCTGCGTTTCGCAATCCGTGAAGCGCGGTACCGTTGGC 1155
QY 1417 GCGGGTGTATCAACAAATCATTTAGTAA 1446
Db 1156 GCGGGCGTTTGTCTAAAGTTCTGGGCTAA 1185

RESULT 5

US-09-815-242-9866
; Sequence 9866, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9866
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Salmonella typhi

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1230)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1230)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-242-9866

Query Match 30.1%; Score 505.6; DB 10; Length 1230;
Best Local Similarity 65.1%; Pred. No. 8e-127;
Matches 801; Conservative 0; Mismatches 385; Indels 45; Gaps 2;

QY 216 CCGCGAGGGGCAAGTTCGAGCGCACCAACACACGCTCAACATAGGCACCATCGGCCA 275
Db 45 CCGTGTCTAAAGAAAATTTGAACGTACAAACCGCACGTTACGTTCGGCACCATCGGCCA 104
QY 276 TGTGACCAACGGAAGACACACTCTCACCGCGCGGTACCACTGCTGCTGCTGCTGCTG 335
Db 105 CGTTGACCAACGTTAAACACTTACCTGACCGCTGCAATACCACTGCTGCTGCTGCTG 164
QY 336 TGGCAGCGCGCTAAGAAGTACGAGAGTACGACCGCGCGCGCGCGCGCGCGCGCGCG 395
Db 165 CGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 224
QY 396 TATCACCATCAACACCGCGCGCGCTGAGTACGAGACCGAGACCGCGCGCGCGCGCG 455
Db 225 TATCACCATCAACACTTCTACGTTGAATACGATACCGCGCGCGCGCGCGCGCGCG 284
QY 456 CGACTGCCCGCGCGCGCGCGCGCTATGTCAAGAAATATGATCACCAGCGCTGCGCAG 515
Db 285 AGACTGCCCGCGCGCGCGCGCGCTATGTAAAAAATGATCACCAGTGTGCTGCGCAG 344
QY 516 CGGTGCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
Db 345 TGGCGCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
QY 576 CTTCTCGCCAAAGCAAGTCGGTGTTCGCCAAGATCGTTGCTTCCTCAACAAGAGCAT 635
Db 405 CTTGCTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464
QY 636 GGTGACGACGAGGAGTCTCTGAGCTGCTGAGCTCGTTCGAGTTCGAGGTCGCGAGCT 695
Db 465 GGTGATGACGAAGAGTCTCTGGAATGTTGAAATGGAAGTTCGTGAATTCGTCTCA 524
QY 696 CTACGAGTACGACGCGCGAGAGTACCAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
Db 525 GTACGACTTCCGCGCGCGAGCACACGCGCATCTGCTGCTGCTGCTGCTGCTGCTG 584
QY 756 GGTCTCATGTTCAACCCCTGCTTGAAGCGCGCGCGCATGAGTGGTGGTGGTGGT 815
Db 585 -----AGGCGCGAGAGTGGGAGCGGAAATCAT 614
QY 816 CTCGTTGTTGATAAAGTGGATTCTTATTCAGTTCGCGCAGAGCGCAGCTGACCTCCC 875
Db 615 CGAAGTGGCTGCTTCTCTGATTCCTTACATCCCGGAACAGAGCGTGGATGACAAG 674
QY 876 GTTCTGTGCTGCTGTTGAAGATGCTCTTCCATACCGGTCGTGCTGCTGCTGCTGCT 935
Db 675 GTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
QY 936 CCGTATAGAGCGTGGCAGCGTCAAGATTGGTGACAGTTCGATATCGTCGGAATCCG 995
Db 735 TCGTGTAGAAGCGGTATCATCAAGTGGGGAAGAGTGAATTCGTTGTTGTTATCAAGA 794
QY 996 CACCGGGAAGTGCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1055
Db 795 GACTCAGAAGTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
QY 1056 GCGCGGAGACAAATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
Db 855 TCGTGTGAGAAGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914
QY 1116 CATGCTGCTGGCAAGCGCTGCTCTATCACACCGCACCAAGTTTGAGGCTGTTGCT 1175

5

RESULT 9

```

US-09-815-242-7318
; Sequence 7318, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7318
; LENGTH: 1208
; TYPE: DNA
; ORGANISM: Helicobacter pylori

```

```

FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1200)
US-09-815-242-7318

Query Match      25.9%   Score 436; DB 10; Length 1200;
Best Local Similarity 61.3%; Pred. No. 5.6e-108;
Matches 753; Conservative 0; Mismatches 445; Indels 30; Gaps 2;

QY 219 GCGAGGCGCAAGTTTCGAGCGCCACCAACACACAGTCAACATAGGACCACTCGGCCATGT 278
Db 3 GCGCAAAAGAAAGTTTAAACAGAACTAAGCGCATGTTAAATATTGAAACCACTTGGCATGT 62
QY 279 CGACACCGAAGAACCACTCTCACCGCGCGCTCACCATGCTGCTCGCCCTCCGTCGGTGG 338
Db 63 AGACCATGGTAAAGCACTTTGAGTGCAGCGATTTTCAGCGGTGCTTTCTTTGAAAGTCT 122
QY 339 CAGCGCGCTAAGAGTAGACGAGATCGACGCGCGCCCGCCGAGGAGCGCGCGCGGTAT 398
Db 123 TCCAGAAATGAAGACTATGATATATGATTAACGCCCTGAAGAAAAGAAAGAGGAT 182
QY 399 CACATCAACACCGCCCGCTCGAGTAGACGAGACCGCCCACTACGACACAGCTCGA 458
Db 183 CACTATCTACTCTCTCATTTGAATGATGAGACTGAAACACAGACACTATGCGCATGTGA 242
QY 459 CTGCCCCGCGCAGCGGACTATGTCAAGAAATATGATCACCGCGCTGCGCAGATGGAGCG 518
Db 243 TTGCCAGACACGCTGACTATGTAAACAAATGATCACCGGTGCGGCGCAATGGAGCG 302
QY 519 TGCCATCTCTGCTGATCGGTGCGGCGGCGCCATGCGGAGACCAACAGACACATCT 578
Db 303 AGCGATTTTGGTTGTTCTGCGAGCTGATGGCCCTATGCTCAAACTAGGAGCATATCT 362
QY 579 CTTCCCAAGCAAGTCGTTGTTCCCAAGATGTTGCTTCTCTCAACAGAGGACATGGT 638
Db 363 ATTGCTCTGCTCAAGTAGCGGTGCTCATCATGTTGTTTCTTAAACAAACAGACATGT 422
QY 639 CGACACGAGGAGCTGCTCGAGCTGCTGAGCTCGAGTCCGCGAGCTGCTCAGCACTA 698
Db 423 AGATGACCAAGAAATGTTAGAACTGTAGAAATGGAAGTGGCGAAATGTTGAGCGGTA 482
QY 699 CGAGTACGACGCGGACGACGCTACCAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
Db 483 TGAATTTCTGCGGATGACACTCTCTATCTGAGCGGTTTCAGCTTTAAGAGCTTTAGAAG- 541
QY 759 TCTCATGTCACCTCTGCTTGAAGCGCGGCGGACGATGAGTGGTCTGCTACATCTCTC 818
Db 542 -----AAGCAAGGCTGGTAAATGTTGGGTGAATGGGTGAAAAAGTGCFTAA 587
QY 819 GTTGGTTGATAAGTGGATTCCTATATTCAGTCCCGCAGAGGAGAGCTGACCTCCCGTT 878
Db 588 ACTTATGCTGAAGTGGATGCTATATCTCTACTCCAGAAAGAGACACTGAAAAACTTT 647
QY 879 CTTGCTCGCTGTTGAAGATGCTTCTCCATCACCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 938
Db 648 CTTGATGCGGTTGAAGATGCTTCTCTATTCGCGGTAGAGGAGCTGCTGCTGCTGCTGCTG 707
QY 939 TATAGAGCTGCGCAGCTGCTGAGATGTTGACACAGCTCCATATCGTCGGAATCCGGACAC 998
Db 708 GATTGAAGAGCGGTGGTGAAGTAGGCGATGAAGTGAATGCTGTTGTTGATCAGACCTAC 767
QY 999 CCGGAAGTGCAGGCTGCTGCTGTTGAGATGTTCCAGAAAGACCACTGATGATGCTATGGC 1058
Db 768 ACAAAAAGCACTGTAAACCGGTGTAGAAATGTTTAGGAAAGAGTGGAAAGGTTGAGC 827
QY 1059 CGGAGCAATGTTGGGCTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
Db 828 CGCGCATATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
QY 1119 GGTGCTGCAAGGCTGCTGCTATCACACCGCAGCAAGTTGAGGCTGTTGCTATGT 1178
Db 888 GGTTCCTATGCAAAACAGGTTCTCATCTCCGACCAAGAAATTTGAGGAGAAATTTATGT 947

```

```

RESULT 10
US-08-831-310-3
; Sequence 3, Application US/08831310
; Patent No. US2002026035A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold et al.
; TITLE OF INVENTION: Helicobacter GHPO 1360 and
; TITLE OF INVENTION: GHPO 750 Polypeptides and Corresponding Polynucleotides
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,310
; FILING DATE: 01-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,175
; REFERENCE/DOCKET NUMBER: 06132/037001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 118..1314
; OTHER INFORMATION:
;
US-08-831-310-3

Query Match      25.9%   Score 436; DB 8; Length 1448;
Best Local Similarity 61.3%; Pred. No. 6.1e-108;
Matches 753; Conservative 0; Mismatches 445; Indels 30; Gaps 2;

QY 219 GCGAGGCGCAAGTTTCGAGCGCCACCAACACACAGTCAACATAGGACCACTCGGCCATGT 278

```

Db 120 GCAAAAGAAAGTTTAAACAGAACTAAGCCGCATGTTAATATGGAACCATTTGGGCATGT 179
Qy 279 CGACACGAAAGACCACTCTACCGCGCGCTCACCATGGTCTCGCTCCGCTCGGTGG 338
Db 180 AGACCATGTTAAAGAGCTTTGAGTGCAGCATTTTACGCGTGTCTTCTTTGAAGTCT 239
Qy 339 CAGCGCGCTTAAAGATACGAGATGAGCGCGCCCGGAGAGCGCGCCCGGTAT 398
Db 240 TGCAGAAATGAAGACTATGATAATATTGATAACGCCCTGAAGAAAAAAGAGAGGAT 299
Qy 399 CACCATCAACCGCCCGCTGAGTAGAGAGACCGGACCGCACTACGCACACGTCGA 458
Db 300 CACTATCCCTACTTCTCATATTGAATAGAGACTGAAACAGACATATCGCATGTGGA 359
Qy 459 CTGCCCGCCGACCGCACTATGTCAGAAATATGATCACCGCGCTCGCAGATGGACGG 518
Db 360 TTGCCAGGACCGCTGACTATGTAATAACATGATCACCGGTGCGGCGCAATGGACGG 419
Qy 519 TGCATCTCTGCTGATCGGTGCGGACGCGGCCCATCCGCGAGACCAAGAGCACATCCT 578
Db 420 AGCGATTTTGGTTTCTGCGAGCTGATGGCCCTATGCCTCAAACTAGGAGCATATCTT 479
Qy 579 CCTGCCAAGCAAGTCGGTGTCCAGATCGTTGTCTTCTTCTCAACAGAGACATGTT 638
Db 480 ATTGTCTCGTCAAGTAGCGGTGCTCATCTGTTTCTTAAACAAACAGACATGTT 539
Qy 639 CGAGCAGGAGGCTGCTCGAGCTGCTGAGCTCGGAGCTCGCTCAGCAACTA 698
Db 540 AGATGACCAAGAATTTAGAACTTGTAGAAATGGAAGTGGGAATTTGTAGCGGTA 599
Qy 699 CGAGTACAGCGGACGACGCTACCAATTCGTGCTGCTCCGCCCTCAAGCGCTCGAGGC 758
Db 600 TGAATTTCTCGCATGACATCTCTATCGTAGCGGGTTCAGCTTAAAGAGCTTTAGAAG - 658
Qy 759 TCTCATGTCAACCTCGCTTGAAGCGCGGAGATGAGTGGTGGCTACATCTTCTC 818
Db 659 -----AAGAAAGGCTGTAATGTGGGTGAATGGGTGAAAGAGTCTTAA 704
Qy 819 GTTGGTTGATAAGTGGATCTTATATCCAGTCCCGCAGAGGACGACTGACCTCCGTT 878
Db 705 ACTTATGGCTGAAGTGGATGCTATATCCCTACTCCGAGAAAGAGACACTGAAAAACTTT 764
Qy 879 CTTCCTCGCTTCAAGATGCTCTCCATCAGCGGCTGCTGAGTACAGTTCGCCACTGGCG 938
Db 765 CTTCATGCGGTTGAAGATGTTCTCTATTCGGGTAGAGGACTGTGTTACAGTAG 824
Qy 939 TATAGAGCTGGCCCGCTCAAGATTGCTGACAGCTGATATCGTGGAAATCCGGACAC 998
Db 825 GATTGAAGAGCGGTGTGAAGTAGCGGATGAAGTGGAAATCGTTGGTATCAGACCTAC 884
Qy 999 CCGAACTGCACGCTCACTGGTGTGAGATGTTCCAGAGACCATGATGCCATGGC 1058
Db 885 ACAAAAACGACTGAACCGGTGTAGAAATGTTTAGGAAAGAGTTGAAAAAGGTGAAGC 944
Qy 1059 CGGAGCAATGTTGGGTGCTGCTCCGTGTATGTCAGAGGATGACATTAAGAGGCGAT 1118
Db 945 CGGCGATAATGTGGCGTGTCTTTGAGAGAACTAAAAAAGAAAGTAGTGAACCGGAT 1004
Qy 1119 GGTCTGGCAAAAGCTGGCTTATCACACCGCACACCAAGTTTGAAGCTGTTGTATGT 1178
Db 1005 GGTCTATGCAAAACAGTTCTATCACTCCGACAGAAATTTGAGGAGAAATTTATGT 1064
Qy 1179 GCTTAAAGAGAGAGGTTGGCGACACTCACCTTTCTTCCCTGGTTACCGCCACAGTT 1238
Db 1065 CTTTCTTAAAGAGAGGCGGAGACACACTCCATCTTCAACCAATTTACCGCCGCAAT 1124
Qy 1239 CTACATCGGACACTGATGTGACGAGGAGTGTGACTACGATTATGAATCACAGGATGA 1298
Db 1125 CTATGCGGCACACTGATGTGACTGCTCTATCAC-----CCTTCTGA 1169
Qy 1299 GGAGGCGAAGATGTGATGCTGGTACCGGTATCAAAATGATTGTTCACTCATCCAGC 1358
Db 1170 AGCGGTAGAAATGTTATGCTGGCGATATGTGAAATCACTGTAGATTGATTAGCCC 1229

Qy 1359 TGTTCCTTGTGAGCAGGCTATGAGTTTGTCTATCCGTGAGGTGTTAAGACCGTTGGTGC 1418
Db 1230 TGTTCGTTAGAGTTGGGAACATAAATTCGGATTCGTGAAGCGGTAGGACCGTTGGTGC 1289
Qy 1419 CGGTGTCATCAACAAATAATCATTTAGTAA 1446
Db 1290 TGTGTGTGTGAGCAATATTATTGAATAA 1317

RESULT 11

US-09-815-242-6962
; Sequence 6962, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6962
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1185)
US-09-815-242-6962

Query Match 25.0%; Score 420.4; DB 10; Length 1185;
Best Local Similarity 60.8%; Pred. No. 9.3e-104;
Matches 747; Conservative 0; Mismatches 436; Indels 45; Gaps 2;

Qy 219 GCGAGGGGCGAGTTCGAGCGACCAACACACAGTCACATAGGACCATCGCCATGT 278
Db 3 GTCTAAAGAAAATTTGAACGTACAAAACCGCACGTTAAACGTGGTACAAATCGGCCACGT 62
Qy 279 GCACACCGAAAGACCACTCTCACCGCGCTCACCATGTTGCTTCGCTCCGTCGGTGG 338
Db 63 TGACACCGTAAACAACTTTAACAGCAGCAATCACACCGCTATTAGCAAAACACTAGG 122
Qy 339 CAGCGCGCTTAAAGATGACGAGATCGACCGCGCCCGCCGAGGAGCGCCCGGGTAT 398
Db 123 TGTGTACGCGGTGCTATTGACCAAAATTTGATACGCGCCAGAAAAGCGCGGTAT 182
Qy 399 CACCATCAACCGCCGCTCGAGTACGAGACCGGACCGCCGCACTACGACACGTCGA 458
Db 183 TACCATCAACACTTCACACGTTGAATACGATACACCGACTCGCCACTACGACACGTAGA 242


```

QY 459 CTGCCCCGCGCAGCGACTATGTCACGAATATGATCACCGCGCTGCGCAGATGGAGG 518
Db 243 CTCTCGGACACGCGACTATGTTAAATATGATTTACTGTGCGGCACAAATGGATGG 302
QY 519 TGCATCTCTGCTGATCGGTGCGGACGGGCCCTGCGGACAGCAACAAAGAGCACATCCT 578
Db 303 TGCATTTTGTAGTAGTACGACCAACAGATGCTTATGCGACAACTCTGTAACACATCTT 362
QY 579 CTCTGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
Db 363 ATTAGTCCCAAGTAGGTGTTCCATATCATCTGATTTCTTAAACAAATCGCATGGT 422
QY 639 CGACGACGAGGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCT 698
Db 423 AGATGACGAAGATTTATTAGATTTAGTGAATGGAAGTCTGTAACCTTCTATCTCAATA 482
QY 699 CAGTACGACGCGGACGAGTACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db 483 TGACTTCCGAGGTGACGATACCAATCGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
QY 759 TCTCATGCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
Db 540 -----CGGCTAGCAGATTTGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 572
QY 819 GTTGTGTTGATAAGTGGATTCCTATATTCAGTCCCGCAGGAGGAGTACCTCCCGTT 878
Db 573 GTTAGCAACCACTTAGATCTTACATCCAGAACGAGTGGATTTGACCAACCGTT 632
QY 879 CTGCTGCTGCTGTTGAAGTGTCTTCTCATCACCCTGCTGCTGCTGCTGCTGCTGCTGCT 938
Db 633 CTTCTTCCATCGAAGATGTTCTCAATCTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
QY 939 TATAGAGCTGGCAGCGTCAAGATTTGTTGACACAGTCTGATCTGCGAATTCGCGGAC 998
Db 693 TGTAGAACGAGTATTATCCGTACAGGTGATGAAGTAGAATCTCGGTATCAAGATAC 752
QY 999 CCGGAATGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
Db 753 AGCGAAACTACTGTACGGGTGTTGAATTTCCGTTAAATTTACTTGACGAAGTCTGTC 812
QY 1059 CGGACCAATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118
Db 813 AGGTGAACCATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 872
QY 1119 GGTGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178
Db 873 AGTATTAGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
QY 1179 GCTTAAGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1238
Db 933 ATTATCAAGATGAAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 992
QY 1239 CTATGCGGACAACTGATGTGACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
Db 993 CTATTTCCGTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037
QY 1299 GGAGCGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
Db 1038 AGGCGTGAATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1097
QY 1359 TGTGCTGCTGAGCAGGATGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1418
Db 1098 AATTGCGATGATCAAGGTTTACGTTTTCGCAATCCGTTGAGGTTGAGGTTGAGGTTGAGG 1157
QY 1419 CGGTGCTCATCAACAAATCATTTAGTAA 1446
Db 1158 AGGCGTCTCCGAAATCATCAATTA 1185

```

RESULT 12
 us-09-815-242-6972
 ; Sequence 6972, Application US/09815242
 ; Patent No. US20020061569A1

```

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6972
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1185)
; US-09-815-242-6972

```

Query Match 25.0%; Score 420.4; DB 10; Length 1185;
 Best Local Similarity 60.8%; Pred. No. 9.3e-104;
 Matches 747; Conservative 0; Mismatches 436; Indels 45; Gaps 2;

```

QY 219 GCGAGGCGGCAAGTTCGAGCGCACCAACACACAGTCAACATAGGACACCATCGCCATGT 278
Db 3 GTCTAAGAAATTTGAAGGTACAAACCGCACGTAAACGTGGGTACAAATCGGCCACGT 62
QY 279 CGACACGGAAGAACCACTCTCACCGCGCGCTCACCATGGTGTCTCGCTCGCGTGG 338
Db 63 TGACCACGGTAAACAACTTTAACACGCAATCAACAACCGTATTGGCAAAACATTACGG 122
QY 339 CAGCGCGCTAAGAAGTACGAGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGGTAT 398
Db 123 TGTGCGAGCGGTGTCATTTCGACAAATTTGATAACGCGCGCAAGAAAGCGCGGTAT 182
QY 399 CACCATCAACACCGCGCGCTGAGTACGAGACCGCGCGCGCGCGCGCGCGCGCGCGTCA 458
Db 183 TACCATCAACACTTACAGCTTGAATACGATACCGGACTCGCCACTATGCAACAGTAGA 242
QY 459 CTGCGCGCGCGCGCGCGCTATGTCAAGAAATATGATCACCGCGCTGCGCAGATGGACGG 518
Db 243 CTGTCGGGACACGCGCGCTATGTTAAAAATATGATTACTGTTGGGCGCAAAATGGATGG 302
QY 519 TGGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578
Db 303 TGTATTTTGTAGTAGTACGACCAACAGATGCTCTATGCGCGCAAACTCTGTGAACATCTT 362
QY 579 CTTGCGCAAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
Db 363 ATTAGTCCCAAGTAGGTGTTCCATATCATCTGATTTCTTAAACAAATCGCATGGT 422
QY 639 CGACGACGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698

```



```

; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Ge
; STREET: 9410 Key We
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM
; MEDIUM TYPE: Diskette
; COMPUTER: HP Vectra
; OPERATING SYSTEM: M
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION
; NAME: Benson, Bob
; REGISTRATION NUMBER: 307446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-82

Query Match 23.4%; Score 394; DB 7; Length 15598;
Best Local Similarity 57.6%; Pred. No. 4.2e-96;
Matches 774; Conservative 0; Mismatches 525; Indels 45; Gaps 2;

Qy 219 GCGAGGGGCAAGTTCGAGCGCCACCAACACACAGCTCAACATAGGACACCATCGGCCATGT 278
Db 11537 GCGAAAAGAAAAATTCGATCGTTCTAAGAACATATCGCTACTATCGGTACATCGGT 11596
Qy 279 CGACACGGAAAGACCACTCTCACCGCGCGCTCACCATCGTGTCTCGCCCTCCGCTGG 338
Db 11597 TGACCATGGTAAACAACATTAACAGCAGCAATCGCTACTGTATTAGCAAAAAATGGTGA 11656
Qy 339 CAGCGCGCCTAAGAAGTACGACGAGATCGACGCGCGCGCGCGAGAGCGCGCGCGGTAT 398
Db 11657 CTCAGTTGCACAATCATATGACATGATTGACAGCTCCAGAGAAAAAGAACGTGGTAT 11716
Qy 399 CACCATCAACACCGCCACCGTCGAGTACGAGACGAGACCGCGCATACGCCACACGTGGA 458
Db 11717 CACAATCAATCTCTCAGATTTAGTACCAAACTGACAAACGCTACGTACGCTCAGCTTGA 11776
Qy 459 CTGCGCGCGCCACCGCACTATGTCAGAAATATGATCACCGCGCTCGCGAGATGGACGG 518
Db 11777 CTGCGCGAGACACGCTGACATGCTTAAACATGATGATCAGTGGTGTCTCAATGACGG 11836
Qy 519 TGCCATCCTCGTGTATCCGGTCCGCGCGCCATGCCCGACAGCAACAGACATCCT 578
Db 11837 CGGTATCTTAGTATCTGTGCTGACGGTCCAAATGCCCAAACTCGTGAACACATCT 11896
Qy 579 CTGCGCAAGCAAGTCGGTGTCCAGATCGTGTCTTCTTCAACAGAGGACATGGT 638
Db 11897 TTTATCAGCTAACGTGGTGTACCAAGCTTAGTAGTATCTTAAACAAAGTTGACATGGT 11956
Qy 639 CGACGACGAGGCTGCTCGAGCTCGTCGAGCTCGAGGTCGGGAGCTGCTCAGCAACTA 698
Db 11957 TGACGATGAAGAAATTATAGAAATTAGTAGAAATGGAAGTTGCGTACTTATTAACCGAATA 12016

```

```

Qy 699 CGAGTACGACGGGACGACGATACCAATCGTCTGGCTCGCCCTCAAGCGCTCGAGGC 758
Db 12017 TGACTTCCCGAGTACGATGATACCTGTAATCGCTGGTTCAGCATTAAGGCTTTAGA --- 12073
Qy 759 TCTCATGGTCAACCCCTGCTTGAAGCGGGGAGCATGAGTGGGTGCGACTACATCTCTC 818
Db 12074 -----AGCGATGCTCAATACGAAGAAAAATCTTAGA 12106
Qy 819 GTTGGTGTATAAGTGGATTCCTATATTCAGTCCCGAGAGCGAGACTGACCTCCCGTT 878
Db 12107 ATTAATGGAAGCTGTAGATACTTACATTCACACTCCAGAACGTGATTTCTGACAAACAT 12166
Qy 879 CTTGCTCCTGTTGAAGATGCTTCTCCATCACCGGTGCTGTGTACAGTGCACACTGCGCG 938
Db 12167 CATGATGCCAGTTGAGGACGTATCTCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTG 12226
Qy 939 TATAGAGCGTGCACCGCTCAAGATGTTGGTGACACAGTGCATATCTCGGAATCCGGGAC 998
Db 12227 TGTTCAGCGTGTCAAAATCAAAAGTTGGTGAAGAAAGTTGAAATCATCGGTTTACATGAC 12286
Qy 999 CCGGAACCTGCACGCTCACTGCTGTGTGAGATGTTCCAGAAAGACCATGGATGATGCC 1058
Db 12287 ATCTAAAACAACCTGTTACAGGTGTTGAATGTTCCGTAATATTATAGACTACGCTGA 12346
Qy 1059 CGGAGACAATGTTGGGCTGCTGCTCGGTGATGCGAAGAGTACATTTGAAAGAGGAT 1118
Db 12347 TGGTGACAAACATTTGGTGCATTAATACGTGGTGTGCTGCTGAAGACGTACAACTGCTCA 12406
Qy 1119 GGTGCTGCAAAAGCTGCTCTATCACACCGCACACCAAGTTTGAGGCTGTGTTGTATGT 1178
Db 12407 AGTATTAGTCTGCTCTGTTTCAATTTACACCACTACTGCAATTTCAAGCAGAGTATAGCT 12466
Qy 1179 GCTTAAAGAAAGAGGTGGCGGACACTCACCTTTCTTCCCTGGTTACCGCCACAGTT 1238
Db 12467 ATTATCAAAAGACGAAGTGGAGCTCACACTCCATCTCTCTCAAACTATCGTCCACA 12526
Qy 1239 CTACATAGGGCAACTGATGTGACAGGAGTGTGACTACGATTTATGATGACAAAGGATGA 1298
Db 12527 CTATTTCGGTACTACTGACGTAACCTGGTGTGT-----TCACCTTACCAGA 12571
Qy 1299 GGAGCGGAAGATGTCATGCTGCTGGTGACCGTATCAAAATGATTTGCTACGCTCATCCAG 1358
Db 12572 AGGTACTGAAATGGTAAATGCCCTGGTGATACGTTGAAATGACAGTAGAATTAATCGCTCC 12631
Qy 1359 TGTTCCTTGTGAGCAGGTATGAGTTTGTCTATCCGTGAGGTTGGTAAGACCGTTGGTGC 1418
Db 12632 AATCGGATTGAAGACGGTACTCGTTTCTCAATCCGTGAAGGTGGACGTACTGTAGGATC 12691
Qy 1419 CGGTGTCATCAACAAATCAATGAGTAACTGGATATACATATCCACCATGAGAAATTT 1478
Db 12692 AGGCGTTGTTACTGAAATCAATTAATAATTTCTAATTTCTTAGATTTTATATAAAAGAA 12751
Qy 1479 CTTTGTCTTACTCAAGCGACATGCTCCGTAGTTGTTATTATGTTGAGTTTATAGGGGTT 1538
Db 12752 GATCCCTCAATCGAGGGTCTTTTTTAAATGTTGTAATTTTGTAAATGCTATTCGATTTA 12811
Qy 1539 GCTCATGTGCAATTTAGTATGAC 1562
Db 12812 GAAGAACAATAATTGATGAAGAC 12835

```

Search completed: May 1, 2003, 22:09:46
Job time : 1226 secs